

Tilburg University

Misreporting in the Value-Added Tax and the Optimal Enforcement

Hoseini, M.

Publication date:
2014

Document Version
Early version, also known as pre-print

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Hoseini, M. (2014). *Misreporting in the Value-Added Tax and the Optimal Enforcement*. (CentER Discussion Paper; Vol. 2014-061). Economics.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2014-061

**MISREPORTING IN THE VALUE-ADDED TAX AND
THE OPTIMAL ENFORCEMENT**

By

Mohammad Hoseini

6 October, 2014

ISSN 0924-7815
ISSN 2213-9532

Misreporting in the Value-Added Tax and the Optimal Enforcement

Mohammad Hoseini*

October 1, 2014

Abstract

A common fraud by registered traders in the value-added tax system is under-reporting sales and over-reporting purchases. This paper models this problem by linking the level of misreporting to the risk-aversion of taxpayers and the level of transactions with final consumers. In addition, it analyses the enforcement consequences of the new developments in information reporting and electronic invoicing, which enable the tax authority to randomly cross-check the invoices. The results highlight the importance of taxpayer's subjective beliefs in shaping audit policy of the tax authority. The optimal audit rate for risk-neutral firms is an increasing function of transaction with final consumers, but this relationship may turn to be negative for risk-averse taxpayers. Moreover, the optimal level of invoice cross-checking on transactions of each commodity is positively associated with the number of trading firms.

Keywords: Value-added tax, Tax evasion, Information reporting, Predictive analytics

JEL code: H26

1 Introduction

One of the most important concerns of tax administration is choosing a simple and efficient enforcement procedure to reduce the risk of tax fraud and evasion. The value-added tax (VAT) is nowadays the preferred form of indirect tax and is believed to facilitate enforcement through its invoicing system. For tax authorities, each VAT invoice generates a piece of information on a transaction, verifiable in case of inter-firm trade with the corresponding invoice issued by the other party. This type of third-party reporting in the VAT increases the risk of hiding transactions for firms and thus reduces evasion ([Pomeranz, 2013](#)). However, this deterrence mechanism breaks down in transactions with final consumers who do not report the invoice to the government. Consequently, an approach to identify the VAT evasion potential among different taxpayers becomes critical for tax authorities. On the other hand, modern VAT systems involve extensive information reporting to achieve a high level of compliance at a modest cost, but the consequences of such transition is almost neglected in the tax evasion literature. This paper analyzes firm's behavior to evade the VAT and the optimal

*Tilburg University, PO Box 90153 5000 LE Tilburg, The Netherlands, M.Hoseini@uvt.nl. I am grateful to my advisor, Jan Boone, for invaluable guidance and numerous suggestions. I also wish to thank Thorsten Beck, Benedikt Goderis, Manuel Oeschlin, Hans Schumacher, Harrie Verbon and seminar participants in Tübingen University for helpful comments. All errors and omissions are my own.

strategy of the authority in this respect by linking noncompliance with the level of final consumption and risk-aversion of each taxpayers. In addition, it studies the role information reporting, which enables the tax authority to cross-check the VAT invoices, on firm's VAT evasion and discusses its efficient implementation.

Broadly, VAT evasion can be classified into two forms. At the intensive margin, a registered seller gains from under-reporting the sale, while a formal buyer benefits from over-reporting its purchase. I use the generic term *misreporting* to refer to noncompliance by registered traders. In comparison, *informality* represents the extensive margin, i.e. evasion by firms working in the shadow economy and failing to register with tax agencies. In the high-income countries, the loss in VAT revenue due to misreporting is much larger than failure to register. For instance, in 2001-02, non-compliance by registered traders in UK resulted in loss of £6.7-9.75 billion compared to £400-500 million loss from traders not registering for the VAT (Keen and Smith, 2006). However, in developing countries both frauds seems to be extensive, where according to Schneider et al. (2011) the informal sector comprises around 40% of GDP on average (ranging up to 70%). The extensive margin of VAT evasion is studied in Hoseini (2014) in detail, and the focus of this paper is on modeling misreporting problem in the VAT system and how the government can reduce it.

Figure 1 illustrates VAT collection efficiency defined as $VAT\ revenue / (VAT\ rate \times Value-Added)$ versus the share of final consumption in total sales among different activities in UK, where the bulk of loss in the VAT collection is due to misreporting fraud.¹ Unlike business-to-business transactions, no third-party reported invoice exists in business-to-household sales and thus they are less risky for VAT misreporting. The figure justifies this by showing a negative relationship between the two variables in the sense that sectors with more inter-firm transactions (fewer sales to final consumers) have more VAT collection efficiency. In this paper, I pinpoint how the difference in the risk of misreporting affects the optimal enforcement policy of the government.

The model of the paper is based on the standard theoretical model for analyzing tax evasion. The tax authority seeks to minimize VAT evasion given the tax base and the cost of audit. The tax base is composed of firms heterogeneous in VAT obligation, risk-aversion and the type of customers. There are two enforcement instruments available for the authority: (i) audit staff visit the enterprise and check the transactions (visiting audit); (ii) a number of reported business-to-business transactions are cross-checked for correspondence with the other party (random invoice cross-checking). The second

¹According to UK National Audit Office (2004) and UK Government (2013), around one third of firms in UK under-declare their VAT obligation resulting around 8-10% loss in the government revenue and this number is a significant fraction of the 12.2% VAT gap in UK over 2005-2011.

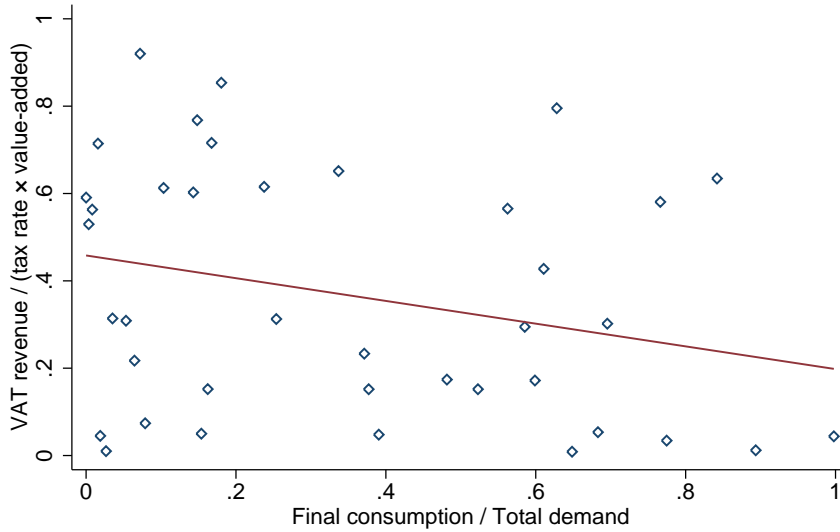


Figure 1: VAT collection efficiency ratio and final consumption share among different UK activities (2006-07). For data details see Appendix B.

enforcement method is much easier for the authority, but it is limited to inter-firm transactions and needs infrastructure on information reporting.

The results of the model intuitively imply that the optimal visiting audit rate negatively depends on the taxpayer's risk-aversion with respect to tax evasion gamble. However, the relationship between optimal visiting audit rate and the level of final consumption of a firm is more complicated. Since transactions with final consumers are unverifiable, on the one hand, they persuade firms to evade more, on the other hand, they reduce the expected return of a visiting audit. For a risk-neutral taxpayer, the optimal audit rate is an increasing function of sales to final consumers. In comparison, more unverifiable transactions does not sharply change the evasion of a risk-averse taxpayer, but it makes the detection of the fraud harder. As a result, for a very risk-averse taxpayer, the optimal audit rate is first increasing but after some point decreasing in sales to final consumers. In practice, large enterprises are more risk-averse and unlikely to engage in gross evasion, such as making uninvoiced transactions. Among other reasons, their accounting systems would not permit this and having numerous employees makes their collusion to unreport tax fragile (Kleven et al., 2009; Tait, 1991). Therefore, VAT collections improves if audit is directed to medium and small enterprises which are more risky taxpayers. This result is at odds with the conventional perception in many countries that the audit should primarily be devoted to larger taxpayers having higher additional tax assessment per visiting audit. In addition, the model show the optimal invoice cross-checking rate among different commodities is positively associated with the number of firms in the market of that commodity. If the cross-checking policy is not chosen optimally, it motivates firms to shift their misreporting to the

commodity that generates the lowest risk.

This paper contributes to several aspects of tax enforcement literature. Most studies about the optimal design of the VAT presume that it is collected costlessly, however, VAT evasion may create a critical impact on its optimality versus other taxes like tariff ([Emran and Stiglitz, 2005](#)). To the best of my knowledge, this is the first paper modeling VAT misreporting with an important policy implications for tax authorities in both developed and developing countries.

Despite the limited research on the VAT specific aspects of tax enforcement, the broad concept of tax compliance has attracted a lot of attention in the literature. As a general classification, [Slemrod and Gillitzer \(2014\)](#) categorize tax evasion models based on their assumption about ethical factors, though the two groups are not mutually exclusive. The ‘deterrence’ models presume the actions of taxpayers are not set by morality or social norms, but are based on the possibility of audit and punishment. This branch of literature can be traced back to the seminal formulation of [Allingham and Sandmo \(1972\)](#), assuming that, to understate their income, risk-averse taxpayers are constrained by a possible penalty and the expected payoff determines the level of evasion. In comparison, ‘non-deterrence’ models focus on the behavioral aspects of tax evasion arguing that [Allingham and Sandmo \(1972\)](#) cannot fully explain the compliance rates, especially in the developed world (for a survey see [Hashimzade et al., 2013](#)). This growing branch of the literature mainly studies the effect of social norms and ethical parameters such as regret, shame, or delight at cheating on the behavior of taxpayers. [Andreoni et al. \(1998\)](#) indicate that, by tax evasion, people may fear social stigma or damage to reputation suggesting that factors such as a moral obligation to be truthful, or the social consequences of being a known cheater, may add further compliance incentives that are not accounted for in the standard models. In the paper, using a simple deterrence framework, I explain how the subjective beliefs of a taxpayer are reflected in his cost function and analyze their role in shaping the optimal audit policy.

The paper also contributes to the literature emphasizing the importance of different forms of third party reporting in the tax enforcement. [Johns and Slemrod \(2010\)](#) show that in the U.S., only 1 percent of wages and salaries is misreported, but in a sharp contrast, an estimated 57 percent of self-employment business income, which is based on self-assessment, is not reported. In a field experiment on the individual income tax in Denmark, [Kleven et al. \(2011\)](#) show that income tax evasion is low, except for the fraction that is self-reported. Directly relevant to this paper, [Pomeranz \(2013\)](#) uses a randomized experiment on Chilean firms to investigate the effect of the third party reported paper trails in the VAT on the tax compliance. She shows that in transactions with final

consumers, in which the VAT paper trail is absent, the response of firms to an audit letter message is stronger, consistent with the stylized fact provided in Figure 1.

The results of the paper also add to the recent and growing literature investigating the effect of new developments in information reporting systems on tax compliance. As a general definition, information reporting are the requirements that certain transactions causing tax obligation be reported by the third party to the tax authority (Shaw et al., 2010). By adding information reporting into the standard tax enforcement framework, Paramonova (2014) models how tax authority can affect the accuracy of information about taxpayers and determines the optimal tax audit rule for a given information accuracy. This paper also models the consequence of information reporting system, but in a VAT setting, and characterizes the optimal way of using such a system.

Finally, this paper contributes to the literature on predictive analytics and subjective audits in tax enforcement. In the deterrence models, the audit is usually assumed to be random; however, the majority of audits conducted in practice are risk-based and not completely random. Some papers model the optimal audit strategy as a function of reported income and the dynamics of taxpayer's behavior (e.g. Andreoni et al., 1998; Reinganum and Wilde, 1985). Another important factor determining tax fraud is taxpayer's risk attitude. The effect of risk aversion on tax evasion is traced back to Allingham and Sandmo (1972) who predict, *ceteris paribus*, risk-averse individuals evade less than others. Considering all of these factors, some papers suggest simulation-based approaches to predict the compliance incentive of each taxpayer, and consequently, its optimal audit rate (e.g. Engel and Hines Jr, 1999; Hashimzade et al., 2014). Analyzing the risk of taxpayers, helps overcome the asymmetric information problem the tax authority faces, and I assume using predictive analytics methods, the authority clusters the firms based on their risk-aversion and final consumption. The results of model show that such analysis increases collection efficiency and is an essential element of optimal VAT enforcement.

After this introduction, this paper provides some empirical evidence using two datasets from manufacturing sector of India and examines the factors explaining VAT misreporting. The findings indicate that the chance of VAT misreporting of a commodity is increasing in its final consumption, number of traders, and the size of producers. Next, in section 3, I explain the practical VAT enforcement issues and model the behavior of a single firm when it decides between misreporting the transactions with other firms and final consumers. Sections 4 and 5, respectively, study the optimal policies in the presence and absence of an information reporting system which enables the authority to perform random invoice cross-checking. Section 6 concludes.

2 Preliminary evidence

The previous evidence suggests a negative relationship between final consumption and VAT compliance of a sector. [Tait \(1991\)](#) indicates that the mark up ratio of the VAT, defined as reported sales divided by reported purchases in a period, is less in retail and wholesale sectors. In a randomized experiment, [Pomeranz \(2013\)](#) observes that firms respond more to an increase in audit probability on transactions with final consumers, where there is no paper trail, suggesting more ex-ante evasion on these transactions. In this part, I provide some evidence about the effect of final consumption of a commodity on its VAT misreporting from India. According to [Hoseini \(2014\)](#), among Indian small services sector enterprises that are above VAT registration threshold, just about 12% are registered with tax authorities between 2001 and 2006, suggesting high informality rate. Moreover, the below estimation shows around 31% VAT misreporting among manufacturing units. Therefore, as a developing country, both informality and misreporting are high in Indian VAT system.

The data used in the paper is from the manufacturing sector of India in 2005-06, given the two rich surveys conducted then by the Ministry of Statistic and Programme Implementation, Government of India. The main data is drawn from Annual Survey of Industries (ASI) by Industrial Statistic Wing of Central Statistical Office. This survey contains the representative sample of 37,055 formal manufacturing establishments with more than 10 employees and has detailed information on various aspects of firms' activities. The second data source is an enterprise survey by National Sample Survey office (NSS) which includes a representative sample of 80,637 small manufacturing units that are not covered by ASI. In both dataset, proper weight showing how many enterprises the sample represents is provided.² In the ASI, the firm-level information about the details of quantity, value of sales, and excise duties of each product and by-product is available. This information enables us to measure tax evasion by the level of underpaying the tax obligation on sales. Nevertheless, the statistical agency is not authorized to reveal the identity of the enterprise.³

There are two types of indirect taxes on manufacturing establishments in India: Excise duty by the central government and sales tax by state governments. In this paper, I just focus on excise duties to avoid state level differences in tax rates and VAT adoption. In India, excise duties (Cent-VAT) are levied on all manufacturing products with the same ad-valorem rate per commodity all over the

²The data description and the questionnaires are available online at micro data archive of MOSPI <http://mail.mospi.gov.in/>.

³In general, measuring evasion and misreporting is not simple since firms are reluctant to give such information. Nevertheless, ASI is conducted independent of tax purposes and no information about the identity of the enterprise is supposed to go to the other agencies. As it mentioned in the manual of ASI, according to the Collection of Statistics Act, violation of any of the confidentiality and secrecy of the information by statistics officer or field staff is prosecuted by or with the consent of appropriate agency.

country. In addition, under central VAT system, the assessee himself determines the duty liability on the goods and clears the goods, thus the audit method is self-assessment without invoice reporting.

In order to measure misreporting per commodity, in the first step, we need the tax rates. The rate of excise duties is changing across commodities and overtime and unfortunately there is no reliable information about it. To overcome this, instead of de jure tax rate –which may not be immediately applied and be complied by firms– I measure the de facto rate by finding the mode of the rate paid by firms in ASI, after dropping the zero payments.⁴ Having tax rate per commodity, the overall misreporting in manufacturing sector can be estimated by the average of *excise evasion*/(*tax rate* × *production*), weighted by the level of production, which equals 31.1%.

For testing the effect of commodity characteristics on its tax underreporting, we also need reliable data on the intermediate and final consumption and also the number of sellers and buyers. Both ASI and NSS have commodity-level information about the input materials and outputs of each firm including quantity, value, unit and the commodity 5-digit code. I estimate the measures by combining and aggregating the data using sample weights. First, I compute the sum of sales and purchases of each commodity by manufacturing firms as well as the number of its sellers and buyers. Next, I estimate two indexes for final consumption of each commodity as⁵

$$\text{final consumption per firm} = \log \left(1 + \frac{\text{total sales} - \text{purchases by other firms}}{\text{total number of sellers}} \right) \quad (1)$$

$$\text{final consumption share} = \log \left(1 + \frac{\text{total sales} - \text{purchases by other firms}}{\text{total sales}} \right) \quad (2)$$

Regarding VAT underreporting, first I compute the overall excise payment and sales of each commodity in ASI, and then I construct two measures:

$$\text{evasion per firm} = \log \left(\frac{\text{tax rate} \times \text{value of sales} - \text{excise payment}}{\text{total number of sellers}} \right) \quad (3)$$

$$\text{evasion to tax base} = \log \left(\frac{\text{tax rate} \times \text{value of sales} - \text{excise payment}}{\text{tax rate} \times \text{value of sales}} \right) \quad (4)$$

Table 1 shows the standardized coefficients of OLS estimation of evasion on commodity characteristics. The unit of observation is commodity and to control for the similarity between different

⁴If the hypothesis of less compliance of sectors with high final consumption is true, we expect to observe lower de facto rate in those sectors. Therefore, the estimations are likely to underestimate the effect of final consumption on misreporting.

⁵If the purchases are higher than the sales, I assume the product is completely used as intermediate input and the difference is provided by imports. That is the reason of adding 1 in the logarithm. Unfortunately, I do not find any corresponding commodity-level import and export tables of India in 2005-06 to take them into account.

goods, I cluster the standard errors at 2-digit Indian ASICC commodity code. The results indicate that both absolute and relative evasions are positively affected by final consumption share of the commodity. One standard deviation increase in the commodity's final consumption per firm increases evasion per firm and evasion to tax base by 3.9 and 7.9 percent respectively. These numbers are 2.6 and 5.7 percent for the effect of final consumption share. In addition, more sellers and buyers both significantly increase evasion of a commodity, which can reflect the higher, the number of units in a sector, the lower, the chance of audit and the risk of detection. The average production per firm increases evasion of the commodity too, but it is a concave function since the square term has a negative coefficient. This is consistent with [Tait \(1991\)](#) who indicates although misreporting is positively associated with the size, but larger taxpayers do not violate VAT proportional to their production and marginal evasion is decreasing in firm size. In other words, the degree of risk-aversion with respect to tax evasion is positively associated with the firm's size such that evasion of big firms, in relative terms, is smaller than small firms. Finally, while the effect of tax rate on absolute evasion is positive, it is negatively associated with relative evasion. One simple explanation is that, in addition to misreporting, higher tax rate increases total tax base –the numerator of the relative index– and thus has two opposing effects in the latter case.

	evasion per firm		evasion to tax base	
	(1)	(2)	(3)	(4)
final consumption per firm	0.039*** (4.46)		0.079*** (4.32)	
final consumption share		0.026** (2.41)		0.057** (2.55)
log of number of sellers	0.103*** (5.07)	0.105*** (5.21)	0.262*** (6.08)	0.265*** (6.13)
log of number of buyers	0.030*** (3.12)	0.028** (2.66)	0.064*** (3.32)	0.062*** (2.94)
log of production per firm	1.117*** (19.26)	1.124*** (18.76)	0.524*** (4.07)	0.540*** (4.06)
squared log of production per firm	-0.205*** (-3.94)	-0.205*** (-3.83)	-0.506*** (-4.15)	-0.507*** (-4.05)
rate of excise duties	0.159*** (8.39)	0.159*** (8.42)	-0.140*** (-5.91)	-0.140*** (-5.87)
observations	2379	2379	2379	2379

Table 1: The effect of commodity characteristics on its overall misreporting –The unit of observation is commodity. Standardized coefficients and t statistics (in parentheses) are reported. Standard errors are clustered at 2-digit commodity code. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Given the preliminary evidences, I move forward in three steps. First, I present some background

information about the practical issues of VAT audit which lays the foundation of the model. Then I model firm's behavior and government policy when random cross-checking is possible. Afterwards, I extend the model to a case that there is no possibility of cross-checking to see how it affects the outcomes.

3 VAT audit in practice and the basis of the model

In this part, I describe the key aspect of VAT administration in different countries and then base them to characterize the theoretical model. In a VAT system, firms are required to charge tax on their output and in return deduct taxes paid on inputs from their VAT bill. The input credit is given against VAT payment as a refund and thus just VAT-payers can obtain it. The implementation of VAT in almost all countries, except Japan⁶, is based on invoice-credit form, which is the focus throughout the paper. In this method, registered sellers issue an invoice corresponding to the VAT charged on sales to each customer, who if registered, can use the invoice to get refund on inputs. Like any other tax, VAT is vulnerable to evasion and the governments must choose an auditing strategy. According to [Ebrill et al. \(2001\)](#), in general, there are three audit methods for the VAT:

Simple self-assessment without invoice reporting– Each enterprise calculates its own tax liability (usually per month or quarter) and sends the aggregate tax return forms to the authority. It does not have to send the invoices, but has to keep them for some years and is subject to a possible audit by the tax authority. During the audit, the authority checks the book of accounts and extensively cross-checks all of the invoices with the corresponding reports of the firm's suppliers and business clients. The auditors also use other possible information like bank accounts to find out the violation in other transactions. In 2001, around half of countries with the VAT apply this method of VAT administration ([Ebrill et al., 2001](#)).

Self-assessment with invoice reporting– Firms calculate their own tax liability, but also send additional documents such as invoices to the tax authority. The authority then can audit the firm in person or just by cross-checking the invoices. This method is applied in many countries and is growing because of the new developments in information technology.

Direct government audit– The taxpayer files a return in the tax office, then the authorities audit and assess the tax obligation of the firm. The method is not common nowadays and broadly was conducted in 1990s in some former Soviet Union states.

⁶In Japan, each trader is taxed on the difference between sales and purchases.

There is a trade-off between enforcement method and compliance cost of taxpayers. The advocates of the first method argue that it has the least compliance cost and is more efficient. [Ebrill et al. \(2001\)](#) indicate that self-assessment procedures with complex requirements place excessive compliance costs on taxpayers and can be a serious impediment to detection of delinquent taxpayers. Without invoice reporting, however, cross-checking is not possible. The early attempt of heavy cross-checking was done in South-Korea in 1980s. Overall, it was unsuccessful due to complexity of the work and the lack of IT technology. At the same time, some countries like Bolivia and Chile conducted cross-checking on a random basis, but the processes were done manually by tax auditing staff and was very time-consuming ([Tait, 1991](#)). Nevertheless, the development of information technology nowadays has increased tax enforcement performance considerably and e-auditing is now growing all over the world. With the assistance of electronic invoicing and data mining –a methodology to identify specific information from rough data via computing technology– tax authorities can collect the third party information much easier to perform data matching and finding evasion cases ([Wu et al., 2012](#)). Currently, many countries⁷ are adopting integrated electronic invoicing system which enables the government to randomly cross-check the invoices. Dealing with the huge volume of invoices might be impossible and even if the list of all unmatched transactions is available by data mining methods, identifying the fraudulent party –seller or buyer– and the arrangements of proper penalty need auditing staff. Therefore, in reality, tax authorities are able to investigate a fraction of the suspicious transactions for further auditing because of limited staff resources and have to prioritize between different commodities or sectors. Nevertheless, if such a system is developed, its variable cost for cross-checking a firm is much less than the cost of a visiting audit for the tax authority.

In any types of assessment, if a fraud is detected, the fraudulent firm has to pay the misreported tax plus an extra penalty. [Tait \(1991\)](#) indicates that when a taxpayer misreports a small amount, the purpose of the penalty is to dissuade him so that he does not repeat the violation. But when the fraud goes beyond the violation and falls into the realm of crime, harsh penalties, including jail sentences, may apply. In practice, the level of penalty is different across countries. For instance, in UK it is changing from 20% to 100% of the fraud based on its magnitude. In some countries the fraudulent has to pay from 2 to 10 times of the misreported tax (like Argentina and Bolivia). Nevertheless, almost in all countries large scale frauds leads to closing of the business and also years of imprisonment. Hence, in practice the penalty is an increasing and convex function of the level of

⁷some examples are Brazil, Bolivia, Bulgaria, Chile, China, Cote d'Ivoire, Indonesia, Poland and South Korea.

evasion.

3.1 The basic assumptions of the model

In order to model misreporting and optimal enforcement in the VAT, I follow the standard model of tax evasion and extend it by differentiating between two types of transactions. Consider an economy comprising firms heterogeneous in risk-aversion and type and level of tax obligation. Based on the type of activity, each firm makes two types of transactions: with other formal businesses (b), with final consumers (c), imposing y_b and y_c as the VAT obligation. As a result, each firm decides about two types of evasion e_b and e_c based on the nature of the obligation. The difference between b and c comes from their transparency to the tax authority. The authority can realize the exact amount of y_b –and thus e_b – by cross-checking the invoices, but y_c has no corresponding third-party reported information and small amount of e_c may be unrealizable. Detection of e_c is possible only by visiting audits when the authority checks information such as bank accounts, total turnover, size, location etc. The probability that the audit detects e_c depends on the relative extent of the fraud. If the evasion comprises a very small fraction of the tax obligation, the detection would be very hard, but the probability increases when the relative extent of fraud increases. Therefore, I assume the probability that the authority detects the fraud on sale to final consumers in a visiting audit is e_c/y_c .

There are two enforcement methods for the tax authority. The first one involves random cross-checking and needs infrastructure on information reporting. The second one is self-assessment without invoice reporting. In the first method, the authority has two separate tools for each type of evasion. It randomly cross-checks a share of inter-firm transactions to detect e_b , but e_c can be detected only by visiting audit. In the simple self-assessment system, no random invoice cross-checking is possible and if a visiting audit takes place, the auditors thoroughly checks all VAT invoices for correspondence, as well as other information for estimating y_c . As a results, the fraud in transactions with formal firms (e_b) is for certain detected, but the probability that the authority detects the fraud on sale to final consumers is e_c/y_c .

If a fraud is detected, a penalty is applied by the authority which is always greater than the amount of evasion. If the detected evasion is near zero the penalty approaches to the principal of unpaid obligation. Based on the facts explained above, I assume $\theta(e)$, an increasing and convex function of the detected evasion e with the following characteristics, determines the cost of fraud for the firm:

$$\theta(0) \geq 0, \quad \text{and} \quad \forall e > 0: \theta(e) > 0, \theta'(e) > 1, \theta''(e) > 0, \theta'''(e) \geq 0 \quad (5)$$

These assumptions about the cost of detection just reflect the financial costs for a risk-neutral taxpayer. In practice, other factors like risk aversion and social norms can increase the convexity of cost function for some taxpayers. In below, I take these measures into account when discussing the optimal visiting audit policy. Prior to that, the results does not depend on the functional form of $\theta(e)$.

In this paper, although the decision of firm for evasion (e_c and e_b) depends on the enforcement strategy of the tax authority, production and tax obligation y_b and y_c are held fixed. This is the common assumption in models for analysis of tax evasion, like the seminal formulation of [Allingham and Sandmo \(1972\)](#) which has been the dominating theoretical model in this literature. This assumption is specially plausible for the VAT evasion, since the VAT is essentially a tax on final consumption and it does not distort the profit maximization of firms. Hence, one can reasonably assume that the decisions about production and VAT evasion are orthogonal. Besides, the aim of this paper is characterizing the optimal audit strategy for the tax authority and the determinants of the tax base is not the subject of discussion.⁸ In order to find the optimal visiting audit policy, the paper seeks to find the audit rate as a function of the estimated tax obligation on sales to final consumers and risk-aversion of the firm. Moreover, as it will be discussed in section [4.1.1](#), the optimal policy for invoice cross-checking is independent of tax obligation and depends on the number of firms in the economy.

In the following, for better mathematical tractability, I first analyze the firm's decision and the optimal policy, when an invoice cross-checking system is available for the tax authority, and then study the model in the absence of such system (simple self-assessment).

4 The model with random invoice cross-checking

In this part, I model the consequences of having an integrated invoice system that enables the government for random cross-checking. As mentioned above, such a system is now more feasible with the advancements in information technology and is growing all over the world. In general, there are two types of invoices in the VAT system: (1) *sale invoice* shows the tax payment by the seller, (2) *purchase invoice* is used by the buyer to get credit on inputs. Sales can be understated in the form of not reporting a number of sale invoices by the seller, in comparison the buyer can

⁸Minimizing evasion per se is an important objective from normative point of view, since it reduces the tax inequality between compliant and non-compliant taxpayers. Moreover, unequal tax burden intensifies tax distortion at the aggregate level.

create some fake purchase invoices to overstate the credit due on its input. Therefore, when the government accumulates all reported invoices in one place, some sale invoices are missing and some extra purchase invoices are generated by the buyers. For performing cross-checking, after random selection of n purchase invoices, the government cross-checks them with the corresponding sale invoice. The number n is a policy variable determined by the cost of cross-checking and total number of purchase invoices N in the economy. The government seeks to find the optimal allocation of n across different intermediate goods.

The strategy of the tax authority after drawing and cross-checking n purchase invoices is finding the firms that made the violation. I assume, after detection of a violation, the authority cross-checks all other invoices of the suspicious firm, reimburses the principal tax, and charges an extra penalty. I also assume that, in a firm's view, invoice cross-checking is conducted independent of visiting audits and it optimizes the evasion on inter-business transactions e_b and final consumer sales e_c in two separate problems. This assumption enables me to find the analytic solutions of e_b and e_c by separating the effects of the two enforcement tools. In Appendix A.1, I relax this assumption by assuming the two types of evasion are jointly determined in one optimizations problem and show that in practical circumstances the two problems become very similar.

Before analyzing the general random invoice cross-checking problem, I study a simple market of one intermediate good in which an upstream sector u sells the good to one downstream sector d . The optimization problem of a single firm in sector u (d) is how many sale (purchase) invoices to misreport given the risk of government detection. Each invoice represents a unit of transaction with value α based on the type of the intermediate good.⁹ If a firm misreports k invoices of the intermediate good, its evasion in this market will be $e_b = \alpha k$, and the probability that no misreported invoice is cross-checked becomes $(1 - k/N)^n$ where, N and n are the total number and the number of cross-checked invoices in market of the intermediate good. In practice, the number of reported invoices in a market is so large that a single firm cannot change it and takes it as given,¹⁰ therefore I assume that $k \ll N$ and thus $k/N \approx 0$. As a result, we can use first order approximation and write $(1 - k/N)^n = 1 - kn/N$, which means that the probability of detection of at least one misreported invoice by the government and getting fined is kn/N , where $k = e_b/\alpha$. Then, the optimization

⁹ α shows the average value of one invoice of the good. This assumption is made to distinguish between commodities with high value per unit (e.g. Steel bar) and others (e.g. bread).

¹⁰For instance, around 26 billion VAT invoices are reported in Germany in 2006 (Baldwin, 2007).

problem of a single firm –either in u or d – can be written as

$$\max_{e_b} \quad e_b - \mu \frac{e_b}{\alpha} \theta(e_b) \quad (6)$$

where $\mu = n/N$ is the share of invoices that are cross-checked by the government and the second term represents the expected cost of detection through invoice cross-checking. Then, the FOC becomes

$$\theta(e_b) + e_b \theta'(e_b) = \frac{\alpha}{\mu} \quad (7)$$

Because θ is a strictly increasing and convex function, one can easily show that e_b decreases when n/N goes up. Therefore, when the authority cross-checks a larger share of invoices, inter-firm misreporting reduces.

On the other hand, the authority conducts visiting audits aiming at detection of evasion on unverifiable transactions y_c . As a consequence, the firm faces another optimization problem on e_c due to the expected cost of such audits

$$\max_{e_c} \quad e_c - \lambda \frac{e_c}{y_c} \theta(e_c) \quad (8)$$

where λ stands for the share of firms that are randomly audited by the tax authority and e_c/y_c is the probability of detection of the unverifiable fraud. Then the choice of firm for e_c can be found from the FOC

$$\theta(e_c) + e_c \theta'(e_c) = \frac{y_c}{\lambda} \quad (9)$$

Proposition 1. *In the presence of invoice cross-checking, total evasion $e_b + e_c$ is decreasing in visiting audit rate λ and invoice cross-checking rate μ and increasing in VAT obligation on final consumption y_c .*

Proof. Appendix [A.2](#). □

Proposition 1 indicates that if the government utilizes the cross-checking technology there is a negative relationship between final consumption and VAT compliance. Moreover, total evasion is decreasing in visiting audit and cross-checking rates. In the next step, I discuss the optimal enforcement policy for cross-checking and visiting audits.

4.1 Optimal enforcement

Assume we have an economy with heterogeneous firms and I intermediate goods. The tax authority is not fully informed about the heterogeneity of the firm, but using predictive analytics can categorize them into S different clusters based on the type of activity, the estimated amount of VAT obligation, and risk-aversion.¹¹ Each intermediate good is produced by one activity and is used as an input by a number of activities. The policy vector of the government includes $\mu_i = n_i/N_i$ which is the ratio of transactions of each commodity i to be cross-checked and λ_s which is the audit rate in each cluster s . The aim of this section is finding the optimal allocation of μ_i and λ_s across commodities and clusters based on aggregate and firm-level factors. In the following, first I find the optimal policy for the cross-checking rate μ_i , then explain the role of subjective measures in evasion and determine the optimal visiting audit rate.

4.1.1 Optimal invoice cross-checking

One of the policy instruments of the government is the cross-checking share for each commodity. With the advanced information reporting system, the government receives the invoice of each transaction and decides on the share of input credit invoices of each intermediate good to be drawn for cross-checking (μ_i). In this case, the policy maker's question is finding the optimal allocation of μ_1, \dots, μ_I . If the number of sellers and buyers of intermediate good i are M_u^i and M_d^i and the buyers have no other input –thus they can just over-report the purchases of i –, the optimization problem for the government will be

$$\min_{\mu_i} (M_u^i + M_d^i)e_b^i + \delta\mu_i N_i \quad (10)$$

where δ is the variable cost of cross-checking one invoice and finding its violating firm, \bar{N}_i is the real number of firm-to-firm transaction invoices, and $N_i = \bar{N}_i + M_d k_i$ is total number of purchase invoices.

Proposition 2. *The optimal share of invoice cross-checking μ_i is positively associated with $\frac{M_u^i + M_d^i}{\bar{N}_i}$. In addition, if the variable cost of cross-checking δ is small enough, then at the government's optimum*

$$\mu_i = \min \left[1, \sqrt{\frac{M_u^i + M_d^i}{\delta \bar{N}_i} \alpha_i} \right] \quad (11)$$

Proof. Appendix A.3. □

¹¹The factors to help predict the heterogeneity are discussed in section 4.1.3.

Proposition 2 states the condition for optimal cross-checking rate of a commodity, when the number of traders and invoices are known. Because all firms trading the commodity i have the same chance of getting caught, and in case of detection of one misreported invoice, all verifiable transactions are cross-checked, the incentive for misreporting is the same among them. Therefore, the higher the number of firms, the more the overall misreporting. In addition, if the variable cost of cross-checking δ diminishes by new technologies, the audit rate can be estimated by (11). This estimation is independent of the form of cost function, because small δ results in small e_b , and near zero $\theta(e_b) \approx e_b$. If δ becomes very small, then the optimal decision of the government is cross-checking all transactions of the commodity i .

In the model, the upstream firms produce just one product and their misreporting is limited to invoices of good i , but in (10) and Proposition 2, I assume that the buyers also has just one intermediate good as input. In a general framework, the customer firms may have more than one intermediate input and, as a result, decide about the evasion of each input separately. Before studying the general case, assume that there is one downstream which buys two differentiated inputs from two upstream activities 1 and 2 and it can over-report the purchases of either 1 or 2.

The government draws the shares μ_1 and μ_2 input credit invoices of the two commodities and all firms take the total number of invoices N_1 and N_2 as given. In this case, the optimization problem for the upstream firms are the same as (6), but the probability of not being detected for the downstream is $(1 - k_1/N_1)^{n_1}(1 - k_2/N_2)^{n_2}$. Since N_1, N_2 are large it can be approximated by $1 - \mu_1 k_1 - \mu_2 k_2$. Because $k_i = e_{bi}/\mu_i$, the optimization problem of downstream becomes

$$\max_{e_{b1}^d, e_{b2}^d \geq 0} e_{b1}^d + e_{b2}^d - \left(\frac{\mu_1}{\alpha_1} e_{b1}^d + \frac{\mu_2}{\alpha_2} e_{b2}^d \right) \theta(e_{b1}^d + e_{b2}^d) \quad (12)$$

which gives the FOCs as

$$1 - \left(\frac{\mu_1}{\alpha_1} e_{b1}^d + \frac{\mu_2}{\alpha_2} e_{b2}^d \right) \theta'(e_{b1}^d + e_{b2}^d) - \frac{\mu_1}{\alpha_1} \theta(e_{b1}^d + e_{b2}^d) \leq 0 \quad (13)$$

$$1 - \left(\frac{\mu_1}{\alpha_1} e_{b1}^d + \frac{\mu_2}{\alpha_2} e_{b2}^d \right) \theta'(e_{b1}^d + e_{b2}^d) - \frac{\mu_2}{\alpha_2} \theta(e_{b1}^d + e_{b2}^d) \leq 0 \quad (14)$$

To have nonnegative e_{b1}^d and e_{b2}^d both FOC cannot hold when $\frac{\mu_1}{\alpha_1} \neq \frac{\mu_2}{\alpha_2}$. In this case, at the optimum, at least one of the variables is zero and based on the amount of μ_1/α_1 and μ_2/α_2 we have three possibilities:

1. if $\frac{\mu_1}{\alpha_1} < \frac{\mu_2}{\alpha_2}$, we have $e_{b2}^d = 0$ and $e_{b1}^d = e_{b1}$

2. if $\frac{\mu_1}{\alpha_1} > \frac{\mu_2}{\alpha_2}$, we have $e_{b1}^d = 0$ and $e_{b2}^d = e_{b2}$
3. if $\frac{\mu_1}{\alpha_1} = \frac{\mu_2}{\alpha_2}$, the two FOC yield the same result $e_{b1} = e_{b2}$ which means d is indifferent between the two commodities and we have $e_{b1}^d + e_{b1}^d = e_{b1}$

Next, the government's optimization for μ_1 and μ_2 is obtained by solving

$$\min_{\mu_1, \mu_2} M_1^u e_{b1} + M_2^u e_{b2} + M^d \max[e_{b1}, e_{b2}] + \delta(\mu_1 N_1 + \mu_2 N_2) \quad (15)$$

According to Proposition 2, at the government's optimum, μ_i is an increasing function of the number of evading firms. Besides, Proposition 1 indicates that for each commodity i , e_b is a strictly decreasing function of μ_i . Therefore, when the tax authority sets its policy from Proposition 2, we can write the equilibrium misreporting of intermediate good i as

$$e_{bi} = h_i(M_i) \quad (16)$$

where $h_i(\cdot)$ is a decreasing function of M_i , the total number of firms that misreport the commodity i . If $\exists i \in \{1, 2\}$ such that $h_i(M_i^u + M^d) > h_j(M_j^u)$, then μ_1 and μ_2 are determined independently and all e_b^d is on commodity i . However, if $h_i(M_i^u + M^d) > h_j(M_j^u)$ never holds, the optimal strategy is setting $\frac{\mu_1}{\alpha_1} = \frac{\mu_2}{\alpha_2}$ such that $e_{b1} = e_{b2}$. In the general setting, the formulation of optimal policy is obtained in the same way.

Proposition 3. *Consider a downstream activity that uses J different intermediate goods as input. Derive the minimum number $j \in \{1, \dots, J\}$ that can separate the goods into two mutually exclusive sets P_j and Q_j with j and $J - j$ elements respectively, in the sense that $\forall p \in P_j, q \in Q_j$, we have $h_p(M_p^u + \frac{1}{j}M^d) > h_q(M_q^u)$ (if $Q = \emptyset$ put $j = J$). Then, the optimal policy $\forall p_1, p_2 \in P_j$ is*

$$\frac{\mu_{p1}}{\alpha_{p1}} = \frac{\mu_{p2}}{\alpha_{p2}}$$

In addition, $\forall q \in Q_j$, the downstream misreporting for commodity q is zero and the optimal policy is determined from Proposition 2, given the number of other traders.

Proof. Appendix A.4. □

Proposition 3 provides the optimal policy of the government when a downstream activity has multiple inputs. At the optimal equilibrium, these conditions hold in all downstream activities that

have multiple intermediate commodities as inputs. The optimal decision of a firm in such activities is over-reporting the input(s) that has the lowest cross-checking probability. Therefore, the government should allocate cross-checking rates across commodities such that a single commodity does not attract a large share of firms for misreporting. This suggests that commodities that are the input of a lot of other activities and have a lot of traders should be under more cross-checking than others. Here, for simplicity, I assumed that each activity produces just one commodity, but may use a number of inputs. In a general setting, if suppliers also have a number of products to sell, the authority should take into account that they can also switch between different products for evasion in the form of under-reporting the sales. Then, Proposition 3 should be extended to the case that sellers can also choose which output to under-report.

4.1.2 Objective versus subjective costs of fraud

So far, I discussed the role of sectoral characteristics on firm's evasion and the assumptions about cost of detection in (5) just reflect the financial costs for a risk-neutral taxpayer. This is a reasonable assumption to formulate the optimal cross-checking policy where the tax authority deals with invoices not firms. For finding optimal visiting audit rate, however, the type of the taxpayer also plays an important role in determining the expected return of the auditing case. In practice, taxpayers usually do not have similar impressions about the cost of evasion and their subjective views about risk, reputation, and social norms differentiates their attitudes toward tax fraud (Andreoni et al., 1998). For some people, conviction for tax evasion is unimaginable due to fear of social stigma or damage of reputation, but a habitual criminal may think about it just as a worst case scenario. In the model, the subjective characteristics of a firm are reflected in its cost function $\theta(\cdot)$, in the sense that the convexity of the function is higher for risk-averse taxpayers. In order to specify a measure for the subjective costs, first, I find evasion as a function of factors that are exogenous for firms.

Lemma 1. *In the presence of invoice cross-checking, we can write $e_c = g(\frac{y_c}{\lambda})$, where $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a strictly increasing and concave function and $\forall x \geq 0$, $0 < g'(x) \leq \frac{1}{2}$.*

Proof. Appendix A.5. □

According to Section 3, the cost of fraud is always a convex function and this makes e_c a strictly concave function of y_c/λ . Therefore, taxpayers are heterogeneous in two dimensions: One is y_c which reflects the possibility of evasion; another is $g(\cdot)$ which reflects the subjective costs of each taxpayer. So far, I assumed the same $\theta(\cdot)$ –and thus $g(\cdot)$ – for all taxpayers, but hereafter, I add the possibility

of different functional forms to capture psychological and cultural aspects of taxpayers such as risk aversion, tax morale, patriotism, guilt, and shame. The curvature of $\theta(\cdot)$ and consequently $g(\cdot)$ help to characterize these subjective beliefs for each taxpayer. In order to measure the curvature, I use the conventional index of relative risk aversion (RRA).¹² Specifically, consider e_c as the utility of the firm and y_c/λ as asset –which here is the possibility of evasion– then $\gamma(x) = -xg''(x)/g'(x)$ is the relative risk aversion of the taxpayer toward VAT fraud: The more curved the function $g(\cdot)$ is, the lower will be its certainty equivalent of a risky bundle. In the next part, I use this concept to determine the optimal visiting audit rate for the government.

4.1.3 Optimal audit rate and taxpayers' subjective beliefs

In order to find the optimal visiting audit λ in each cluster, the authority seeks to minimize evasion in sales to final consumers given visiting audit is costly. Therefore, the optimization problem is

$$\min_{\lambda_s} e_c^s + \eta \lambda_s \quad (17)$$

where λ_s is the rate of visiting audit in cluster s , η is the cost of auditing one firm,¹³ and e_c^s is the estimation of authority from the evasion in cluster s . Then, FOC of (17) can be written as

$$\frac{\partial e_c^s}{\partial \lambda_s} = -\eta \quad (18)$$

Thus, the audit rate should be adjusted at the level that the marginal reduction in the evasion of a firm is equal to the (shadow) cost of auditing. This means that at the optimum marginal reduction in the evasion is the same in all clusters.

One important issue for the policymakers is how to overcome information asymmetries to estimate e_c^s . Based on Lemma 1, e_c depends on y_c and $g(\cdot)$, therefore to find e_c^s , the tax authority needs an estimation of those parameters for each firm. The size of sales to final consumer is easier to estimate and can be found by checking the information on bank accounts and location as well as the type of activity (for instance the bulk of costumers of a retailer or a barber shop are final consumers). To estimate the heterogeneity of taxpayers in terms of risk-aversion and other subjective factors, a growing literature proposes the use of predictive analytics (Hashimzade et al., 2014). This type of

¹²Slemrod and Yitzhaki (2002) indicate that tax evasion is akin to the choice of how much to gamble. Each unit of misreporting the VAT offers a payoff but may lead to a penalty.

¹³An equivalent assumption about cost of auditing is assuming the tax agency is able to audit a limited number of firms and η is the Lagrange multiplier or the shadow cost of a single audit.

analysis provides a set of tools to use historical data to predict future outcomes and create a consistent risk ranking of taxpayers. There are a number of factors that help the authority to estimate the subjective cost of taxpayers. According to the evidence in section 2, evasion is a concave function of firm's size suggesting a negative relationship between size and relative risk aversion. The bigger firms usually have numerous employees and often hire external accountants that might mistakenly report the evasion or make a rational whistle-blowing (Kleven et al., 2009). Therefore, the collusion for hiding the evasion breaks easier in a big firm rather than a small one with limited workers, resulting more risk-aversion for bigger firms. The second factor to estimate the cost function is the age of the firm (Feinstein, 1991). Often, younger firms are more likely to evade tax, because at the early stages the enterprise faces more financial constraints and moreover may not be precisely aware about tax rules. The ownership of the firm is another factor to predict the cost of evasion. The entrepreneurial enterprises, run by the owner, are more likely to evade tax rather than managerial enterprises that have an external manager (Egger et al., 2014). Gender and education of the taxpayer has been revealed to be the other significant determinants of tax evasion. Kastlunger et al. (2010) show that the tax compliance of females are normally higher than males. Witte and Woodbury (1985) find a negative association between the general education level of taxpayers and tax evasion. The occupational choice, social networks, past behavior and compliance reputation of taxpayers are other factors revealing the risk behavior of firms (Hashimzade et al., 2014). Define y_c^s and $g_s(\cdot)$ as the estimated value of tax obligation on sales to final consumers and the functional form of $g(\cdot)$ in cluster s . Then, the following proposition shows how the level of audit rate depends on subjective characteristics.

Proposition 4. *In the presence of invoice cross-checking, if λ_s is the optimal choice of the government in cluster s , we have*

$$\lambda_s^2 = \frac{y_c^s}{\eta} g_s' \left(\frac{y_c^s}{\lambda_s} \right), \quad \frac{\partial \lambda_s}{\partial y_c^s} = \frac{\lambda_s}{y_c^s} \frac{1 - \gamma_s}{2 - \gamma_s} \quad (19)$$

$$\text{where } \gamma_s = - \frac{\frac{y_c^s}{\lambda_s} g_s'' \left(\frac{y_c^s}{\lambda_s} \right)}{g_s' \left(\frac{y_c^s}{\lambda_s} \right)}.$$

Proof. Appendix A.6. □

Proposition 4 indicates that the relationship between the optimal audit rate and final consumption of a cluster depends on the relative risk-aversion of taxpayer to evade e_c defined as γ_s . When the

relative risk-aversion in increasing in y_c/λ , it means that when the taxpayer experiences an increase in the possibility of evasion, he chooses to decrease the share of evasion in total tax obligation on final consumptions. Figure 2, illustrates the optimal audit rate for two different functional forms of $g(\cdot)$. From (19), we can write

$$\eta\lambda_s = \frac{y_c^s}{\lambda_s} g'(\frac{y_c^s}{\lambda_s}) \quad (20)$$

and therefore graphically find the optimal audit rate in each point. The slopes of the tangent lines reflect $g'(y_c^s/\lambda_s)$ and the distance from the vertical axis is y_c^s/λ_s . Thus, $\eta\lambda_s$, the multiplication of the two, is the height of the triangle they shape. The below curve with an asymptote stands for higher relative risk aversion and at the optimum has smaller audit rate. The horizontal asymptote in $g(\cdot)$ corresponds to a vertical asymptote in the cost function θ . As mentioned in Section 3, many countries have implemented harsh penalties for big tax frauds and from one angle, the asymptote in the cost function can represent the level of evasion in which the defrauder is convicted to long-term imprisonment or life in prison. However, from different viewpoint, it reflects the subjective beliefs of compliant taxpayers who never imagine evasion more than the asymptote due to high risk aversion or fear of social stigma. In the next step, I characterize the relationship between λ_s and y_c^s based on the different amounts of relative risk aversion γ_s .

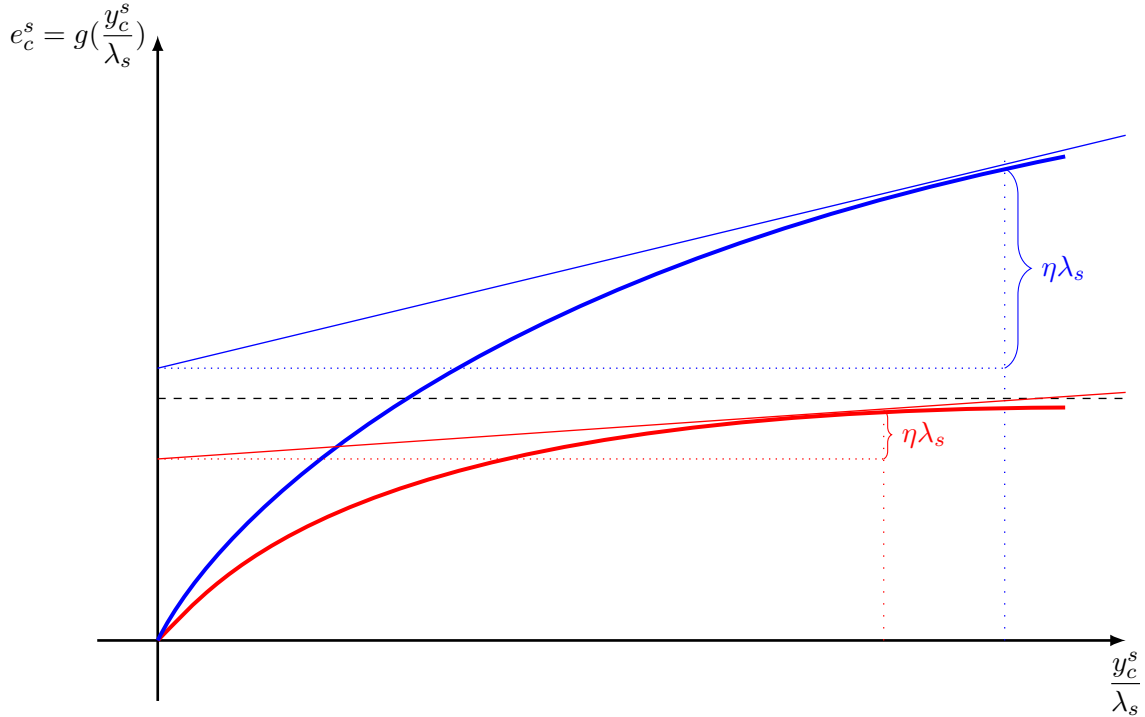


Figure 2: Optimal audit rate for two asymptotic and non-asymptotic cost functions.

Lemma 2. For all $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ that satisfy the conditions of Lemma 1, at the government's

optimum we have

$$\lim_{y_c^s \rightarrow 0} \lambda_s = 0, \quad \lim_{y_c^s \rightarrow 0} \gamma_s = 0, \quad \lim_{y_c^s \rightarrow 0} \frac{\partial y_c^s}{\partial \lambda_s} = +\infty,$$

Proof. Appendix A.7. □

Lemma 2 shows when y_c^s is near zero, λ_s is zero but its derivative is $+\infty$. In this neighborhood the relationship between λ and y_c^s is independent of γ_s . However, for a larger level of y_c , the relationship becomes very sensitive to the relative risk aversion of taxpayers. Using Proposition 3, we can distinguish between four different cases:

1. low risk aversion ($\gamma_s < 1$): λ_s is increasing in y_c^s .
2. medium risk aversion ($\gamma_s = 1$): λ_s is constant in y_c^s .
3. high risk aversion ($1 < \gamma_s < 2$): λ_s is decreasing in y_c^s .
4. very high risk aversion ($\gamma_s \geq 2$): $\lambda_s = 0$ and e_c^s is equal to its horizontal asymptote.

For any form of $g(\cdot)$ that satisfies the conditions of Lemma 1, at $y_c^s = 0$ the relative risk aversion is zero, but it can change at higher level of y_c^s . For better interpretation, I use four different functional forms for $g(\cdot)$ and show how they shape the link between λ_s and y_c^s . The details of calculations are available in the appendix A.8. Figure 3a shows this relationship for $g(\cdot)$ as a n -th root function $g(x) = \frac{1}{2}na(\sqrt[n]{1+x/a}-1)$ where $a > 0, n > 1$. In this example, $\gamma_s = \frac{n-1}{n} \frac{x}{x+a}$ and $\forall x > 0, \gamma_s < 1$. Therefore, λ_s is an increasing function of y_c^s and for large enough amounts of y_c^s , the optimal λ_s has a corner solution equal to one.¹⁴ The second example is when $g(\cdot)$ is a logarithmic function $g(x) = \frac{1}{2}a \ln(1+x/a)$, $a > 0$. Here, $\gamma_s = \frac{x}{x+a}$ and λ_s is increasing in y_c^s , but as y_c^s approaches infinity, $\partial \lambda_s / \partial y_c^s$ approaches zero which means, unlike 3a, λ_s never reaches 1. The other two cases are for taxpayers who are highly risk averse such that $g(\cdot)$ has a horizontal asymptote. In Figure 3c, $g(x) = \frac{a}{2} \frac{x}{a+x}$ and $\gamma_s = \frac{2x}{a+x}$. Therefore, by increasing y_c^s , at first λ_s increases, then reaches a maximum when $\gamma_s = 1$ ($y_c^s = a\lambda_s$), and afterward, where $1 < \gamma_s < 2$, it becomes decreasing. The last case in Figure 3d is for very high risk averse taxpayers where $g(x) = \frac{1}{2n} \left[1 - \left(1 + \frac{x}{a} \right)^{-\frac{1}{n}} \right]$ and $\gamma_s = \frac{(n+1)x}{a+x}$. Similar to Figure 3c, around $y_c^s = 0$, λ_s is increasing, and when $1 < \gamma_s < 2$, it becomes decreasing. When γ_s approaches 2, which means $\frac{y_c^s}{\lambda_s} \geq \frac{2a}{n-1}$, then $\frac{\partial \lambda_s}{\partial y_c^s} = -\infty$ and λ_s drops to zero and remains constant for higher levels of y_c^s . In this case, the amount of evasion stays equal to the horizontal asymptote of $g(\cdot)$ which is equal to $\frac{a}{2n}$. The reason of the negative

¹⁴In reality, because the cost of audit η is large, this hypothetical solution is reached in very large levels of y_c^s that does not exist in practice.

relationship between audit rate and final consumption in the last two examples is the existence of asymptote in the cost function. Basically, no evasion is possible above the asymptote and higher final consumption just makes detection of fraud difficult (lower e_c/y_c), without increasing evasion.

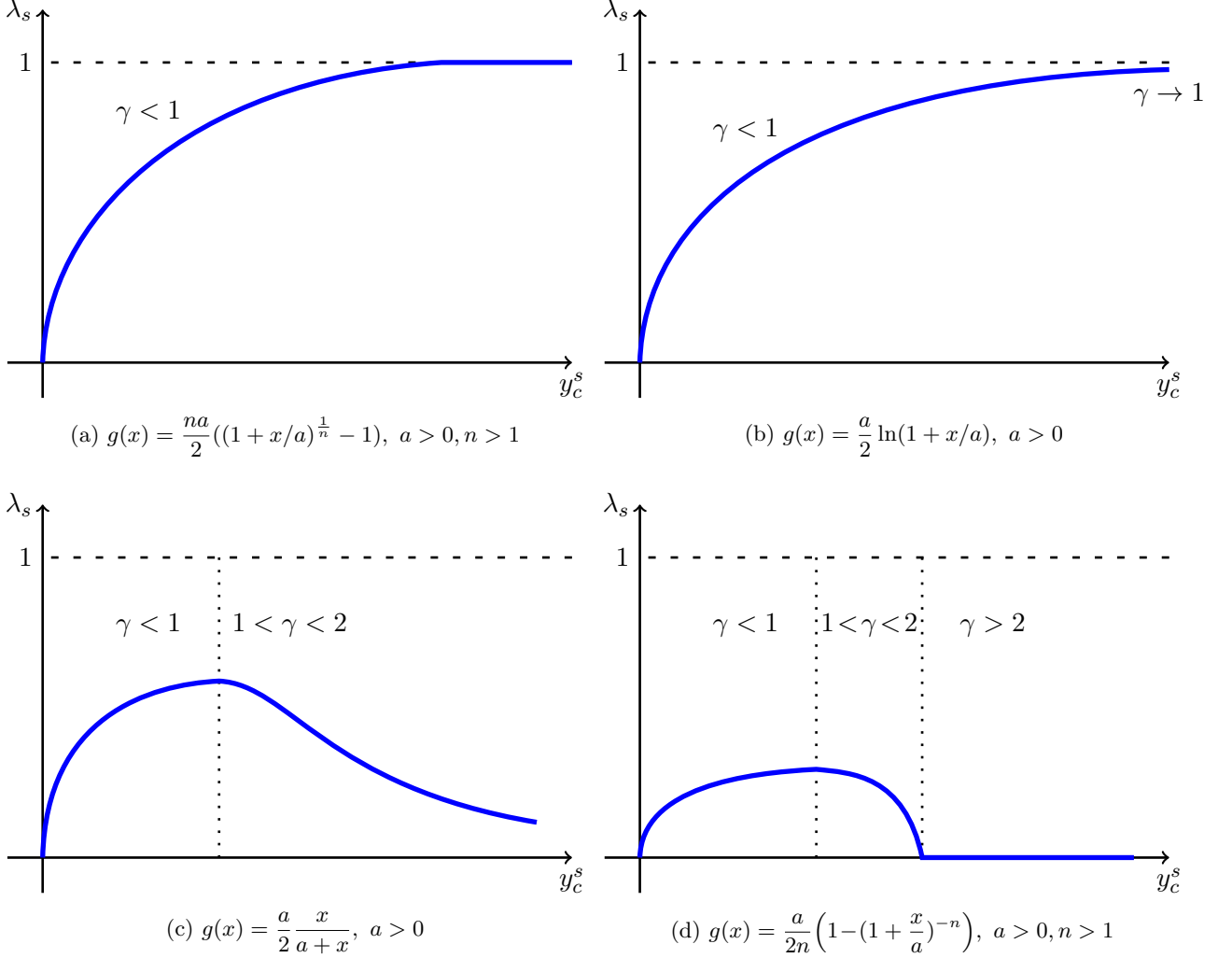


Figure 3: Government audit and firm's final consumption. For risk neutral firms the relationship is always positive, but for very risk averse taxpayer it may turn to be decreasing or become zero.

Figure 3 illustrates the importance of subjective beliefs of tax payers in shaping the optimal visiting audit rate of the tax authority. Therefore, all factors mentioned above can play a role in determining the audit rate of a firm. For instance, large taxpayers are unlikely to engage in gross evasion to avoid hurting their reputation. Therefore, they have higher γ_s than smaller taxpayers. This may lead to less optimal audit rate for them and thus, the total revenue may decline if the audit of medium and small taxpayers is neglected. This is in contrast with the conventional wisdom in many countries to over-allocate the audit staff in larger firms. In addition, the government should collect information about various characteristics of the firm such as age, ownership, gender

and education of the manager to have a better estimation of its risk behavior. As discussed above, a young and entrepreneurial firm with male and uneducated manager tend to have smaller risk-aversion. Therefore, the optimal audit policy for them can be decreasing in size of sales to final consumers. Overall, the results of this section highlight the importance of predictive analytics in determining audit policy.

5 Self-assessment method without invoice cross-checking

In this section, I assume that the VAT administration uses self-assessment without invoice reporting and at each time randomly audits a share of firms. When a visiting audit takes place, all VAT invoices are checked for correspondence and e_b is revealed, but the chance of detecting e_c depends on the relative extent of the fraud and is equal to e_c/y_c . In this setting the optimization problem of the firms can be written as

$$\max_{e_b, e_c \geq 0} e_b + e_c - \lambda \left(\left(1 - \frac{e_c}{y_c}\right) \theta(e_b) + \frac{e_c}{y_c} \theta(e_b + e_c) \right) \quad (21)$$

The FOCs are¹⁵

$$\lambda \theta'(e_b) + \lambda \frac{e_c}{y_c} \left(\theta'(e_b + e_c) - \theta'(e_b) \right) = 1 \quad (22)$$

$$\lambda \frac{e_c}{y_c} \theta'(e_b + e_c) + \frac{\lambda}{y_c} \left(\theta(e_b + e_c) - \theta(e_b) \right) = 1 \quad (23)$$

Condition (22) indicates that, at the optimum, the marginal gain of one addition unit of e_b is equal to the sum of the expected deterministic and stochastic marginal costs of it, which are $\lambda \theta'(e_b)$ and $\lambda e_c/y_c (\theta'(e_b + e_c) - \theta'(e_b))$ respectively. Moreover, according to (23), e_c results in two marginal costs: one is marginal cost of the rise in the punishment which is $\lambda e_c/y_c \theta'(e_b + e_c)$ and the other is the marginal expected cost due to the increase in detection probability of e_c . At the optimum these two costs are equal to the marginal benefit of one additional unit of evasion.

If (21) is a concave function, then the solutions given by FOC are the optimal choice and they maximize the objective function. However, the concavity of (21) near the critical point of FOC depends on y_c , λ and the cost function.

¹⁵Here, I implicitly assumed that the amount of tax obligation on business-to-business transactions y_b is large and imposes no restriction on e_b . If the solution of e_b in (21) is greater than y_b , it has the corner solution $e_b = y_b$ and the optimization is just w.r.t. e_c .

Proposition 5. *In the self-assessment enforcement method, if*

$$\left(\theta'(e_b + e_c) - \theta'(e_b)\right)^2 < \left(2\theta'(e_b + e_c) + e_c\theta''(e_b + e_c)\right) \frac{\theta(e_b + e_c) - \theta(e_b)}{\theta'(e_b)} \theta''(e_b) + 2e_c\theta''(e_b + e_c)\theta'(e_b)$$

then the optimal choices of firm for e_c and e_b are found using (22) and (23). Otherwise, at the optimum, $e_b = 0$ and e_c is found from $e_c\theta'(e_c) + \theta(e_c) = y_c/\lambda$.

Proof. Appendix A.9 □

Proposition 6 states the optimal firm's decision for evasion. When the inequality holds, the objective function is concave, otherwise the FOC results in a saddle point where the objective function has maximum in one direction and minimum in the other. In this case, the best decision of firm is reallocating the evasion from business transactions to sales to final consumers. The saddle point solution happens when $\theta''(e_b)$ is small and y_c –and as a results e_c – is large. For instance, if $\theta''(0) = 0$, then for small enough e_b the inequality becomes $(\theta'(e_c) - 1)^2 < 2e_c\theta''(e_c)$ which does not hold for large enough e_c . (a simple example is $\theta(e) = e^n + e$). In such situations, the firm is better off by being fully compliant for e_b . Nevertheless, if $\theta''(0)$ is large enough, for a simple polynomial or exponential cost functions, one can show that the saddle point never exists. In appendix A.11, I analyze the inequality for different functional forms.

The optimality conditions of the firm, either if there is a maximum or a saddle point, intuitively suggest that higher audit rate increases compliance, but more final consumption reallocates evasion in favor of e_c , increasing the overall amount. Formally, we can show that the results of Proposition 1 are valid in this context.

Proposition 6. *In the self-assessment enforcement method, total evasion is increasing in transactions with final consumers y_c and decreasing in visiting audit rate λ .*

Proof. Appendix A.10. □

Proposition 6 is consistent with the evidence of the positive relationship between final consumption and total evasion of a firm. Therefore, this relationship is independent of the audit method of the government and is valid in the absence of invoice cross-checking.

5.1 Optimal audit policy

In the simple self-assessment method without cross-checking, the optimal audit rate is obtained similar to Section 4.1.3. Here, e_b^s also depends on λ_s and consequently the FOC w.r.t. λ_s can be

written as

$$\frac{\partial(e_c^s + e_b^s)}{\partial\lambda_s} = -\eta \quad (24)$$

In this case, define $e^s = e_b^s + e_c^s$, and from (22) and (23) assume

$$e^s = G\left(\frac{y_c^s}{\lambda_s}, y_c^s\right), \quad G_i = \frac{\partial G(x_1, x_2)}{\partial x_i}, \quad G_{ij} = \frac{\partial^2 G(x_1, x_2)}{\partial x_i \partial x_j} \quad (25)$$

then in the Appendix A.12, I show

$$\lambda_s^2 = \frac{y_c^s}{\eta} G_1, \quad \frac{\partial\lambda_s}{\partial y_c^s} = \frac{\lambda_s}{y_c^s} \frac{1 - \gamma + \zeta}{2 - \gamma} \quad (26)$$

where $\gamma = \frac{y_c^s}{\lambda_s} \frac{G_{11}}{G_1}$ and $\zeta = y_c^s G_{12}/G_1$. In general, finding the boundaries of ζ is cumbersome, however if $0 < \zeta < 1$, the qualitative results in above about the effect of relative risk aversion on the link between λ_s^* and y_c^s are still valid. As a robustness check, I simulate the optimal condition for λ_s using two simple quadratic and asymptotic functional forms. If $\theta(x) = \frac{1}{2}x^2 + x$ and $\theta(x) = \frac{1}{1-x} - 1$ the qualitative relationship between λ_s and y_c^s are similar to Figure 3a and Figure 3c respectively. For the quadratic cost function the audit rate is monotonically increasing in final consumption, but for the asymptotic cost function the relationship is first increasing and then overturns at a specific point.

By comparing self-assessment method with invoice cross-checking, we can simply show that at the same rate of visiting audit, total evasion is less or equal when invoice cross-checking is available (see Appendix A.13). However, a more general question for policymakers is the choice between different enforcement methods. Answering this question requires information about the fixed cost of implementing an integrated invoice system as well as variable costs of cross-checking and visiting audit and is not in the scope of this paper. But intuitively, we can say as new technologies reduce the fixed costs of such system the governments are better off by utilizing it.

6 conclusion

The recent developments in tax enforcement literature highlight the importance of third party reporting and information technology on reducing tax fraud and evasion. Research in this field, however, has tended to focus on direct taxes. This paper models how these two instruments can be optimally utilized in the value-added tax which nowadays emerges as the indirect tax of choice and is adopted

by more than 150 countries.

Using enterprise-level estimation of Indian manufacturing sector, I show that, together with the number and size of traders, final consumption of a commodity increases the intensive margin of VAT fraud (misreporting). This effect can be explained by the invoice system of the VAT which raises the risk of evasion on business-to-business transaction and makes sales to final consumers more attractive for VAT fraud. Using this variation in the detection risk, the model implies that evasion is an increasing function of sales to final consumers and audit rate. It also shows how invoice reporting limits the evasion on inter-firm transactions by enabling the tax authority to do random cross-checking.

In addition to the new framework for analyzing VAT evasion, this study provides important implications for tax authorities. First, it pinpoints how invoice cross-checking can control VAT evasion and presents how to find its optimal rate, based on sector and commodity characteristics such as number of traders and the level of production. Second, the findings stress the importance of final consumption which is a sectoral characteristic, easy to measure for the policymakers. For instance, the visiting audits should be more focused on downstream sectors and services such as retailer shops, hotels, and restaurants. Third, the paper underlines the significance of risk-based visiting audits and the importance of taxpayers' subjective measures in determining their audit rate. Since large firms are more risk-averse, total VAT revenue can diminish if the audit of medium and small taxpayers is neglected.

Although this paper focuses on the VAT, the results can be extended to other type of taxes in which the authority faces information with different levels of accuracy. For instance, withholding income taxes are third-party reported and can be checked through information reporting systems. In comparison, self-employed income tax should be investigated by visiting audits based on the characteristics of the taxpayer.

Despite the paper adds substantially to our understanding on VAT misreporting, a number of potential limitations need to be mentioned. The present study has only investigated one dimension of VAT fraud and tax authorities must think about other potential evasions while implementing their policies. [Keen and Smith \(2006\)](#) outline a list of VAT frauds and some practical solution. As mentioned above, a notable fraud in developing countries is failure to register which is studied in [Hoseini \(2014\)](#) in detail. The VAT also has a big potential for cross-border frauds like missing trader (carousel) fraud in which the importer disappears after the import (for details see [Keen and Smith, 2006](#)).

References

- Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.
- Andreoni, J., Erard, B., and Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, 36(2):pp. 818–860.
- Baldwin, R. (2007). EU VAT fraud. <http://www.voxeu.org/article/vat-fraud-part-1>. VOX column, 14 June 2007.
- Ebrill, L. P., Keen, M., Bodin, J.-P., and Summers, V. P. (2001). *The Modern VAT*, volume 1. International Monetary Fund.
- Egger, P., Erhardt, K., and Keuschnigg, C. (2014). Personal income taxes, corporate prot taxes and the heterogeneous tax sensitivity of firm-level investments. Working paper.
- Emran, M. and Stiglitz, J. (2005). On selective indirect tax reform in developing countries. *Journal of Public Economics*, 89(4):599–623.
- Engel, E. M. and Hines Jr, J. R. (1999). Understanding tax evasion dynamics. Working paper, National Bureau of Economic Research.
- Feinstein, J. S. (1991). An econometric analysis of income tax evasion and its detection. *The RAND Journal of Economics*, pages 14–35.
- Hashimzade, N., Myles, G. D., and Rablen, M. D. (2014). The use of predictive analytics to increase tax compliance. mimeo, presented in CEBID taxation, social norm and compliance conference.
- Hashimzade, N., Myles, G. D., and Tran-Nam, B. (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys*, 27(5):941–977.
- Hoseini, M. (2014). Value-added tax and shadow economy: the role of inter-sectoral linkages. *CentER Discussion Paper Series No.*, 2013(36).
- Johns, A. and Slemrod, J. (2010). The distribution of income tax noncompliance. *National Tax Journal*, 63(3):397–418.
- Kastlunger, B., Dressler, S. G., Kirchler, E., Mittone, L., and Voracek, M. (2010). Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2d: 4d). *Journal of Economic Psychology*, 31(4):542–552.

- Keen, M. and Smith, S. (2006). VAT fraud and evasion: What do we know, and what can be done? *National Tax Journal*, LIX(4).
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica*, 79(3):651–692.
- Kleven, H. J., Kreiner, C. T., and Saez, E. (2009). Why can modern governments tax so much? an agency model of firms as fiscal intermediaries. Working Paper 15218, National Bureau of Economic Research.
- Paramonova, Y. (2014). The optimal information reporting and tax audit policy.
- Pomeranz, D. (2013). No taxation without information: Deterrence and self-enforcement in the value added tax. Working Paper 19199, National Bureau of Economic Research.
- Reinganum, J. F. and Wilde, L. L. (1985). Income tax compliance in a principal-agent framework. *Journal of Public Economics*, 26(1):1–18.
- Schneider, F., Buehn, A., and Montenegro, C. E. (2011). Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. *Handbook on the shadow economy*, pages 9–77.
- Shaw, J., Slemrod, J., and Whiting, J. (2010). Administration and compliance. *Dimensions of Tax Design: the Mirrlees Review*, J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds), Oxford University Press.
- Slemrod, J. and Gillitzer, C. (2014). *Tax Systems*, volume 1. The MIT Press.
- Slemrod, J. and Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. *Handbook of public economics*, 3:1423–1470.
- Tait, A. (1991). *Value-added Tax: Administrative and Policy Issues*. International Monetary Fund.
- UK Government (2013). Measuring tax gaps, 2013 edition. *HM Revenue and Custom, U.K. Government*, page 8.
- UK National Audit Office (2004). Tackling VAT fraud. *Report by the comptroller and auditor general HC 357 Session 2003-2004*.
- Witte, A. D. and Woodbury, D. F. (1985). The effect of tax laws and tax administration on tax compliance: The case of the us individual income tax. *National Tax Journal*, pages 1–13.

Wu, R.-S., Ou, C.-S., ying Lin, H., Chang, S.-I., and Yen, D. C. (2012). Using data mining technique to enhance tax evasion detection performance. *Expert Systems with Applications*, 39(10):8769–8777.

A Mathematical appendix

A.1 Relaxing the assumption of independence between invoice cross-checking and visiting audits

If e_b and e_c are not independent and jointly determined, we can write the optimization problem of a firm as the combination of (6) and (8).

$$\max_{e_b, e_c} e_b + e_c - \lambda \left(\left(1 - \frac{e_c}{y_c}\right) \theta(e_b) + \frac{e_c}{y_c} \theta(e_b + e_c) \right) - \frac{\mu}{\alpha} e_b \theta(e_b) \quad (27)$$

Then the FOC w.r.t e_b becomes

$$\lambda \left(\left(1 - \frac{e_c}{y_c}\right) \theta'(e_b) + \frac{e_b}{y_c} \theta'(e_b + e_c) \right) + \frac{\mu}{\alpha} (e_b \theta'(e_b) + \theta(e_b)) = 1 \quad (28)$$

which means that the optimal e_b in this case is smaller than in (7). The FOC w.r.t. e_c gives

$$e_c \theta'(e_b + e_c) + \theta(e_b + e_c) - \theta(e_b) = \frac{y}{\lambda} \quad (29)$$

and we can simply find the changes in e_c when e_b decreases by implicit differentiation

$$\frac{\partial e_c}{\partial e_b} = - \frac{e_c \theta''(e_b + e_c) + \theta'(e_b + e_c) - \theta'(e_b)}{2\theta'(e_b + e_c) + e_c \theta''(e_b + e_c)} < 0 \quad (30)$$

Also because $(1 + \partial e_c / \partial e_b > 0)$ total evasion $(e_b + e_c)$ becomes smaller. Therefore, by relaxing the assumption of independence, we have a reduction in e_b and an increase in e_c in the sense that total evasion declines.

The assumption about the separability of e_b and e_c in optimization is reasonable because in practical circumstances the two problems lead to close outcomes. As mentioned above, in practice, the variable cost of cross-checking is much less than visiting audit, therefore at the government's optimum μ is much larger than λ when the firm has a high share of business customers. In this case, the answer of (28) for e_b is very close to (7). On the other hand, when y_b is small and y_c is large,

tax authority leaves that commodity out from cross-checking and e_b in (27) is determined by corner solution $e_b = y_c$. In general, as $(1 - \frac{e_c}{y_c})\theta'(e_b) + \frac{e_b}{y_c}\theta'(e_b + e_c) \ll \frac{1}{\lambda}$ or y_b is small the two problems leads to very similar results.

A.2 Proof of Proposition 1

Define $f(e) = e\theta'(e) + \theta(e)$. Then, because $\theta(e)$ is strictly increasing and convex, we have $f'(e) = e\theta''(e) + 2\theta'(e) > 0$. On the other hand

$$f(e_b) = \frac{\alpha}{\mu}, \quad f(e_c) = \frac{y_c}{\lambda} \quad (31)$$

which results

$$\frac{\partial(e_b + e_c)}{\partial y} = \frac{1}{f'(e_c)\lambda} > 0, \quad \frac{\partial(e_b + e_c)}{\partial \lambda} = \frac{-y}{f'(e_c)\lambda^2} < 0, \quad \frac{\partial(e_b + e_c)}{\partial n} = \frac{-\alpha}{f'(e_b)\mu^2} < 0 \quad (32)$$

A.3 Proof of Proposition 2

From (10), we can write the cost function of the government as

$$C_i = (M_i^u + M_i^d)e_b^i + \delta\mu_i(\bar{N}_i + M_i^d \frac{e_b^i}{\alpha_i}) \quad (33)$$

At the government's optimum, the first and second order conditions are $\frac{\partial C_i}{\partial \mu_i} = 0$ and $\frac{\partial^2 C_i}{\partial \mu_i^2} > 0$. By defining

$$A_i = \frac{M_i^u + M_i^d}{\delta\bar{N}_i}, \quad B_i = \frac{M_i^d}{\alpha_i\bar{N}_i}$$

we can write

$$\frac{\partial C_i}{\partial \mu_i} = \frac{1}{\bar{N}_i} \left((A_i + B_i\mu_i) \frac{\partial e_b^i}{\partial \mu_i} + B_i e_b^i + 1 \right) \quad (34)$$

$$\frac{\partial^2 C_i}{\partial \mu_i^2} = \frac{1}{\bar{N}_i} \left((A_i + B_i\mu_i) \frac{\partial^2 e_b^i}{\partial \mu_i^2} + 2B_i \frac{\partial e_b^i}{\partial \mu_i} \right) \quad (35)$$

Therefore, at the optimum we have

$$(A_i + B_i\mu_i^*) \frac{\partial e_b^i(\mu_i^*)}{\partial \mu_i^*} + B_i e_b^i(\mu_i^*) + 1 = 0 \quad (36)$$

where μ_i^* is the optimal decision of the government. Now, we can find the derivative of μ_i^* with respect to A_i

$$\frac{\partial e_b^i}{\partial \mu_i^*} + \frac{\partial \mu_i^*}{\partial A_i} \left((A_i + B_i \mu_i) \frac{\partial^2 e_b^i}{\partial \mu_i^2} + 2B_i \frac{\partial e_b^i}{\partial \mu_i^*} \right) = 0 \quad (37)$$

which means

$$\frac{\partial \mu_i^*}{\partial A_i} = \frac{-\partial e_b^i / \partial \mu_i^*}{(A_i + B_i \mu_i) \frac{\partial^2 e_b^i}{\partial \mu_i^2} + 2B_i \frac{\partial e_b^i}{\partial \mu_i^*}} \quad (38)$$

From the second order condition, the denominator of (38) is positive, thus because $\partial e_b^i / \partial \mu_i < 0$, μ_i is increasing in A_i .

Now, if $\delta \ll (1 + \frac{M_i^u + M_i^d}{N_i})\alpha_i$, then B_i is negligible relative to A_i and we can approximate the first order condition (36) as

$$\frac{\partial e_b^i}{\partial \mu_i} = -\frac{1}{A_i} \quad (39)$$

On the other hand, when δ –cost of an additional cross-checking– is small, the level of evasion is small too and we can approximate $\theta(e) \approx e$ (because $\theta' = 1$ near zero). Then, (7) turns to

$$\frac{\alpha_i}{\mu_i} \approx 2e_b^i \quad \rightarrow \quad \frac{\partial e_b^i}{\partial \mu_i} \approx \frac{-\alpha_i}{2\mu_i^2} \quad (40)$$

As a result, $\frac{\alpha_i A_i}{2}$ is a good approximation of μ_i^2 when cost of cross-checking is low. But at most μ can be equal to one and if δ is very small maybe cross-checking of all invoices are optimal.

A.4 Proof of Proposition 3

If a downstream activity uses J intermediate goods as input and over-reports k_1, \dots, k_J invoices of each input, the probability of not being detected for a firm in that activity is $(1 - k_1/N_1)^{n_1} \dots (1 - k_J/N_J)^{n_J}$ and its approximation will be $1 - \sum_{i=1}^J \mu_i k_i$. As a consequence, the optimization problem of a single firm is

$$\max_{e_{bi}^d \geq 0} \sum_{i=1}^J e_{bi}^d - \sum_{i=1}^J e_{bi}^d \frac{\mu_i}{\alpha_i} \theta \left(\sum_{i=1}^J e_{bi}^d \right) \quad (41)$$

The FOC for intermediate good j becomes

$$1 - \frac{\mu_j}{\alpha_j} \theta \left(\sum_{i=1}^J e_{bi}^d \right) - \sum_{i=1}^J e_{bi}^d \frac{\mu_i}{\alpha_i} \theta' \left(\sum_{i=1}^J e_{bi}^d \right) \leq 0 \quad (42)$$

To have positive e_{b1}^d to e_{bJ}^d all FOC must hold and $\forall i, j : \frac{\mu_i}{\alpha_i} = \frac{\mu_j}{\alpha_j}$. Otherwise, at least one e_{bi}^d is zero. Because each upstream activity produces just one intermediate good, its only misreporting

possibility is on one commodity, but, d has different options for misreporting. Similar to the case with two inputs, if $e_{bi} = \max\{e_{b1}, \dots, e_{bJ}\}$, the downstream firm d is better off by over-reporting just the invoices of commodity i . However, this makes the authority to increase the cross-checking of i , because he knows higher number of firms are misreporting on that commodity. Then, increasing the cross-checking rate in i may make risk of other commodities smaller for d and the downstream firms switch their misreporting to other options.

To find the equilibrium, consider the authority makes the downstream firms indifferent between j commodities, that comprise the elements of a set P_j , for over-reporting their invoices and the rest of $J - j$ commodities build up the set Q_j . If downstream firms have no incentive to evade on $q \in Q_j$, on average, the share $1/j$ of them over-report the transactions of $p \in P_j$. In this case, to find the total number of misreporting firms, the authority adds this number to the number of the upstream firms of each intermediate good –that have just one option for misreporting. Therefore, at the government's optimum, if $\forall q \in Q_j$ downstream evasion is zero, then $\forall p \in P_j$, we have $e_{bp} = h_p(M_p^u + \frac{1}{j}M^d)$. On the other hand, P_j and Q_j characterize a Nash equilibrium, if downstream firms have no incentive to switch their over-reporting from a commodity $p \in P_j$ to another $q \in Q_j$. This means that $\forall q \in Q_j$ and $\forall p \in P_j$, we must have

$$h_p(M_p^u + \frac{1}{j}M^d) > h_q(M_q^u) \quad (43)$$

For $Q_j = \emptyset$ we set $j = J$ and the Nash equilibrium always exists, but (43) can hold for $j < J$ too. The optimal equilibrium for the authority is the one that imposes the lowest number of constraints to its optimization problem

$$\min_{\mu_i} \sum_{i=1}^J M_i^u e_{bi} + M^d \max[e_{b1}, \dots, e_{bJ}] + \delta \sum_{i=1}^J \mu_i N_i \quad (44)$$

Because each additional unit of j adds a new binding constraint in form of $\frac{\mu_i}{\alpha_i} = \frac{\mu_j}{\alpha_j}$ to the optimization problem, the equilibrium with the minimum j is the optimal choice of the authority. After finding the minimum j that holds (43), the optimal policy for all $p \in P_j$ is making the downstream firms indifferent for misreporting which means

$$\forall p_1, p_2 \in P_j : \quad \frac{\mu_{p_1}}{\alpha_{p_1}} = \frac{\mu_{p_2}}{\alpha_{p_2}}$$

On the other hand, for all $q \in Q_j$, the downstream firms never decide to misreport commodity q and the optimal policy is determined independently from Proposition 2 considering no downstream firm

over-reports invoices of commodity q .

A.5 Proof of Lemma 1

Similar to Appendix A.2, define $f(e) = e\theta'(e) + \theta(e)$, then according to (7) and (9) $g(x) = f^{-1}(x)$.

The first and second derivative of f are equal to

$$f'(e) = 2\theta'(e) + e\theta''(e) \quad (45)$$

$$f''(e) = 3\theta''(e) + e\theta'''(e) \quad (46)$$

because $\theta' > 1$ (this comes from the assumption that the penalty is always greater than evasion) and $\theta'' > 0$, we have $f'(e) > 2$. Moreover, according to (5), $\theta'''(e) \geq 0$, and thus we have $f''(e) > 0$. Now, according to inverse function theorem, because $g(f(x)) = x$, we have

$$g'(x) = \frac{1}{f'(x)}, \quad g''(x) = -\frac{f''(x)}{(f'(x))^3} \quad (47)$$

Therefore $0 < g'(x) \leq \frac{1}{2}$ and $\theta'''(e) \geq 0$ is the sufficient condition for $g''(x) < 0$.

A.6 Proof of Proposition 4

First, I prove the positive relationship between optimal λ_s and y_c^s in the presence of cross-checking, which is a simpler case because e_b is given, then I generalize the proof for simple self-assessment case. When cross-checking is possible, from (18), the optimal condition for the audit rate require

$$\frac{\partial e_c^s(\lambda_s(y_c^s), y_c^s)}{\partial \lambda_s} = -\eta \quad (48)$$

Here, firm-level evasion e_c^s is a function of government policy $\lambda_s(y_c^s)$ and exogenous characteristic of industry y_c^s . By implicit differentiation with respect to y_c we can write

$$\frac{\partial^2 e_c^s}{\partial \lambda_s \partial y_c^s} + \frac{\partial^2 e_c^s}{\partial \lambda_s^2} \frac{\partial \lambda_s}{\partial y_c} = 0 \quad (49)$$

therefore,

$$\frac{\partial \lambda_s}{\partial y_c^s} = -\frac{\frac{\partial^2 e_c^s}{\partial \lambda_s \partial y_c^s}}{\frac{\partial^2 e_c^s}{\partial \lambda_s^2}} \quad (50)$$

Based on Lemma 1, we can write $e_c^s = g(\frac{y_c^s}{\lambda_s})$ where $0 < g'(x) < 1/2$. In this case, we have

$$\frac{\partial e_c^s}{\partial \lambda_s} = -\frac{y_c^s}{\lambda_s^2} g'(\frac{y_c^s}{\lambda_s}) \quad (51)$$

By differentiating w.r.t. λ_s and y_c^s , we can write

$$\frac{\partial^2 e_c^s}{\partial \lambda_s \partial y_c^s} = -\frac{1}{\lambda_s^2} g'(\frac{y_c^s}{\lambda_s}) - \frac{y_c^s}{\lambda_s^3} g''(\frac{y_c^s}{\lambda_s}) \quad (52)$$

$$\frac{\partial^2 e_c^s}{\partial \lambda_s^2} = \frac{y_c^s}{\lambda_s} \left(\frac{2}{\lambda_s^2} g'(\frac{y_c^s}{\lambda_s}) + \frac{y_c^s}{\lambda_s^3} g''(\frac{y_c^s}{\lambda_s}) \right) \quad (53)$$

Thus, from (50), we obtain

$$\frac{\partial \lambda_s}{\partial y_c^s} = \frac{\lambda_s}{y_c^s} \left(1 - \frac{1}{2 + \frac{y_c^s}{\lambda_s} g''(\frac{y_c^s}{\lambda_s}) / g'(\frac{y_c^s}{\lambda_s})} \right) \quad (54)$$

As a result, by defining $\gamma = -\frac{y_c^s}{\lambda_s} g''(\frac{y_c^s}{\lambda_s}) / g'(\frac{y_c^s}{\lambda_s})$, we can obtain

$$\frac{\partial \lambda_s}{\partial y_c^s} = \frac{\lambda_s}{y_c^s} \frac{1 - \gamma}{2 - \gamma} \quad (55)$$

A.7 Proof of Lemma 2

From Lemma 1, we know g' is always positive and bounded. Therefore, from Proposition 3,

$$\lim_{y_c^s \rightarrow 0} \lambda^2 = \lim_{y_c^s \rightarrow 0} \frac{y_c^s}{\eta} g'(\frac{y_c^s}{\lambda}) = 0 \quad (56)$$

Similarly, because $g'(0) = 1/2$ we can write

$$\lim_{y_c^s \rightarrow 0} \frac{y_c^s}{\lambda_s} = \lim_{y_c^s \rightarrow 0} \frac{\eta \lambda_s}{g'(y_c^s / \lambda_s)} = 0 \quad (57)$$

On the other hand, from (45), $f''(0)$ and consequently $g''(0)$ are both bounded. Therefore, using (57)

$$\lim_{y_c^s \rightarrow 0} \gamma = \lim_{x \rightarrow 0} \frac{x g''(x)}{g'(x)} = 0 \quad (58)$$

Hence

$$\lim_{y_c^s \rightarrow 0^+} \frac{\partial \lambda_s}{\partial y_c^s} = \lim_{y_c^s \rightarrow 0^+} \frac{\lambda_s}{y_c^s} \frac{1}{2} = +\infty \quad (59)$$

A.8 Finding the derivatives of Figure 3

	(a)	(b)	(c)	(d)
$g(x)$	$\frac{na}{2} \left(\left(1 + \frac{x}{a}\right)^{\frac{1}{n}} - 1 \right)$	$\frac{a}{2} \ln \left(1 + \frac{x}{a}\right)$	$\frac{a}{2} \frac{x}{a+x}$	$\frac{a}{2n} \left(1 - \left(1 + \frac{x}{a}\right)^{-n}\right)$
$g'(x)$	$\frac{1}{2} \left(1 + \frac{x}{a}\right)^{\frac{1}{n}-1}$	$\frac{1}{2} \left(1 + \frac{x}{a}\right)^{-1}$	$\frac{a^2}{2(a+x)^2}$	$\frac{1}{2} \left(1 + \frac{x}{a}\right)^{-n-1}$
$g''(x)$	$\frac{1-n}{2an} \left(1 + \frac{x}{a}\right)^{\frac{1}{n}-2}$	$\frac{-1}{2a} \left(1 + \frac{x}{a}\right)^{-2}$	$\frac{-a^2}{(a+x)^3}$	$-\frac{n+1}{2} \left(1 + \frac{x}{a}\right)^{-n-2}$
$\gamma = \frac{-xg''}{g'}$	$\frac{1-n}{n} \frac{x}{a+x}$	$\frac{x}{a+x}$	$\frac{2x}{a+x}$	$\frac{(n+1)x}{a+x}$

A.9 Proof of Proposition 5

By defining

$$e = e_b + e_c, \quad \theta_e = \theta(e_b + e_c) \quad \theta_b = \theta(e_b)$$

we can obtain an alternative formulation for firm's optimization problem

$$\max_{e \geq e_b \geq 0} E(e, e_b) = e - \lambda \left(\theta_b + \frac{e - e_b}{y_c} (\theta_e - \theta_b) \right) \quad (60)$$

Therefore, we have

$$E_e = \frac{\partial E}{\partial e} = 1 - \frac{\lambda}{y_c} \left(\theta_e - \theta_b + (e - e_b) \theta'_e \right) \quad (61)$$

$$E_b = \frac{\partial E}{\partial e_b} = -\frac{\lambda}{y_c} \left((y_c - e + e_b) \theta'_b + \theta_b - \theta_e \right) \quad (62)$$

$$E_{ee} = \frac{\partial^2 E}{\partial e^2} = -\frac{\lambda}{y_c} \left(2\theta'_e + (e - e_b) \theta''_e \right) \quad (63)$$

$$E_{bb} = \frac{\partial^2 E}{\partial e_b^2} = -\frac{\lambda}{y_c} \left(2\theta'_b + (y_c - e + e_b) \theta''_b \right) \quad (64)$$

$$E_{eb} = \frac{\partial^2 E}{\partial e \partial e_b} = \frac{\lambda}{y_c} \left(\theta'_e + \theta'_b \right) \quad (65)$$

Then, the FOC are $E_b = E_e = 0$, equivalent to (22) and (23). The second order condition of partial derivatives to have a concave objective function requires

$$\Delta = E_{ee}E_{bb} - E_{eb}^2 > 0 \quad (66)$$

Given the partial derivatives, the second order condition $\Delta > 0$ is equivalent to

$$(2\theta'_e + e_c\theta''_e)(2\theta'_b + (y_c - e_c)\theta''_b) - (\theta'_e + \theta'_b)^2 > 0 \quad (67)$$

If we expand (67), we must show

$$2\theta'_e\theta'_b + 2\theta'_e(y_c - e_c)\theta''_b + 2e_c\theta''_e\theta'_b + e_c(y_c - e_c)\theta''_e\theta''_b - \theta_e'^2 - \theta_b'^2 > 0 \quad (68)$$

θ is a convex and increasing function, thus $\theta' > 0$, $\theta'' > 0$ and $\theta'_e > \theta'_b$. In addition, from the FOC (62), $y_c - e_c = (\theta_e - \theta_b)/\theta'_b$. Therefore (67) is equivalent to

$$(\theta'_e - \theta'_b)^2 < (2\theta'_e + e_c\theta''_e)\theta''_b \frac{\theta_e - \theta_b}{\theta'_b} + 2e_c\theta''_e\theta'_b \quad (69)$$

If $\Delta < 0$, then the critical point is a saddle point and we have a corner solution. Since both e_b and e are positive and $e \geq e_b$, the corner solution that maximizes the objective function is either $e_b = 0$ or $e_c = 0$ ($e_b = e$). In each case, we have

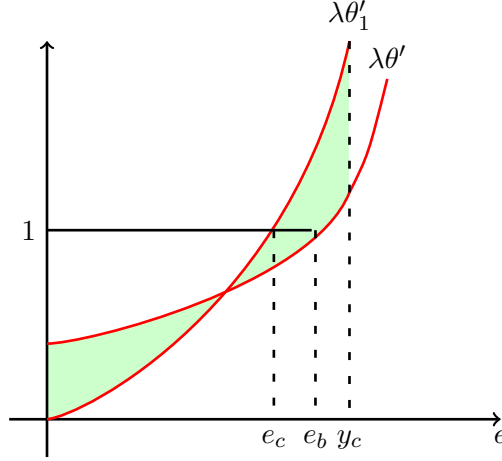
$$e_b = 0 : \quad E(e, 0) = e - \lambda\theta(e) \quad (70)$$

$$e_c = 0 : \quad E(e, e) = e - \lambda \frac{e}{y_c} \theta(e), \quad 0 \leq e \leq y_c \quad (71)$$

By defining $\theta_1(e) = e\theta(e)/y_c$, we can find the optimal decision in each case using FOC

$$e_b = 0 : \quad \theta'(e) = \frac{1}{\lambda}, \quad e_c = 0 : \quad \theta'_1(e) = \frac{\theta(e) + e\theta'(e)}{y_c} = \frac{1}{\lambda} \quad (72)$$

To graphically see $E(e, e) < E(e, 0)$, the below figure illustrates the optimal choice of e_b and e_c when the other is zero. Because $\theta(y_c) = \theta_1(y_c)$ the two shaded areas are equal. The payoff of firm from each corner solution is the area below the horizontal line at 1 and the corresponding curve. It is clearly seen that the area is bigger in case of θ'_1 , and therefore e_c leads to higher payoff



Hence, when $\Delta < 0$, the optimal choice of firm is $e_b = 0$ and e_c is found similar to (9).

A.10 Proof of Proposition 6

If $\Delta < 0$, then the critical point given by FOC is a saddle point which means at the optimum $e_b = 0$ and e_c is found from (9). In this case, the proof is similar to Proposition 1.

If the critical point is not a saddle point and (69) holds, I use the notations of A.9 and in addition I define

$$\dot{e}_b = \frac{\partial e_b}{\partial \lambda}, \quad \dot{e}_c = \frac{\partial e_c}{\partial \lambda}, \quad \dot{e} = \frac{\partial(e_c + e_b)}{\partial \lambda}$$

At the optimum, we have

$$E_b = 0, \quad E_e = 0 \tag{73}$$

Now, if we differentiate the FOCs with respect to λ

$$E_{bb}\dot{e}_b + E_{be}\dot{e} + \frac{\partial E_b}{\partial \lambda} = 0 \tag{74}$$

$$E_{eb}\dot{e}_b + E_{ee}\dot{e} + \frac{\partial E_e}{\partial \lambda} = 0 \tag{75}$$

from (62) and (73), we obtain that $\partial E_b / \partial \lambda = 0$ and $\partial E_e / \partial \lambda = -1/\lambda$, therefore from (74), we can write

$$\dot{e}_b = -\frac{E_{be}}{E_{bb}}\dot{e} \tag{76}$$

By substituting (76) in (75), we obtain

$$\dot{e} = \frac{E_{bb}}{\lambda \Delta} \tag{77}$$

From (64), it turns out that $E_{bb} < 0$. Therefore, because $\Delta > 0$, we have $\dot{e} < 0$. By expanding (77),

we can write the derivative of total evasion w.r.t. λ as

$$\dot{e}_b + \dot{e}_c = \frac{y_c}{\lambda^2} \frac{2\theta'_b + (y_c - e_c)\theta''_b}{(\theta'_e + \theta'_b)^2 - (2\theta'_e + e_c\theta''_e)(2\theta'_b + (y_c - e_c)\theta''_b)} \quad (78)$$

To find the effect of y_c on evasion, define

$$\tilde{e}_b = \frac{\partial e_b}{\partial y_c}, \quad \tilde{e}_c = \frac{\partial e_c}{\partial y_c}, \quad \tilde{e} = \frac{\partial(e_c + e_b)}{\partial y_c}$$

From (61) and (62) we obtain $\partial E_b / \partial y_c = -\lambda\theta'_b / y_c$ and $\partial E_e / \partial y_c = 1 / y_c$. Then, using similar calculation to (76) and (77), we can write

$$\tilde{e}_b = -\frac{E_{be}}{E_{bb}}\tilde{e} + \frac{\lambda\theta'_b}{y_c E_{bb}} \quad (79)$$

$$\tilde{e} = \frac{-1}{y_c \Delta} (E_{bb} + \lambda\theta'_b) \quad (80)$$

By expanding (80), we can write

$$(\tilde{e}_b + \tilde{e}_c) \left(2\theta'_{b+c} + e_c\theta''_{b+c} - \frac{(\theta'_{b+c} + \theta'_b)^2}{2\theta'_b + (y_c - e_c)\theta''_b} \right) = \frac{1}{\lambda} - \frac{\theta'_b}{2\theta'_b + (y_c - e_c)\theta''_b} \quad (81)$$

The left hand side multiplier is equal to Δ / E_{bb} and it is positive. The right hand side is positive if

$$\lambda < 2 + (y_c - e_c)\theta''_b / \theta'_b \quad (82)$$

and it always holds because $0 < \lambda < 1$. Therefore, when $\Delta > 0$ we have $\tilde{e} > 0$.

A.11 Analyzing different functional forms of θ in Proposition 5

Define $f(x) = \theta(e_b)$ and $f(y) = \theta(e_b + e_c)$, then the inequality of Proposition 5 is equivalent to

$$\left(f'(y) - f'(x) \right)^2 < \left(2f'(y) + (y - x)f''(y) \right) f''(x) \frac{f(y) - f(x)}{f'(x)} + 2(y - x)f''(y)f'(x) \quad (83)$$

To check the validity of the inequality for different functional forms, I split the problem to simpler parts and provide example of each case. If $\forall x > 0, f'''(x) \leq 0$, we have

$$(y - x)f''(x) \geq f'(y) - f'(x) \quad (84)$$

Therefore it is sufficient to show that

$$(y-x)f''(x) \leq f''(y) \frac{f(y)-f(x)}{f'(x)} \quad (85)$$

which holds because f is convex and $(y-x)f'(x) \leq f(y)-f(x)$

If $f'''(x) > 0$, the inequality may violate if $f''(0) = 0$ or f is an asymptotic function, therefore I assume in addition to other conditions $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is differentiable on \mathbb{R}^+ and $f''(0) = 1$. Then, we can rewrite (83) as

$$\left(\frac{f'(y)-f'(x)}{y-x} \right)^2 < \left(\frac{2f'(y)}{y-x} + f''(y) \right) \frac{f''(x)}{f'(x)} \frac{f(y)-f(x)}{y-x} + 2 \frac{f''(y)f'(x)}{y-x} \quad (86)$$

As x approaches y , the left and right hand sides approach $f''(x)^2$ and $+\infty$ respectively, and the inequality holds. Therefore, since f and its derivative are both strictly increasing and convex, for a fixed y if the inequality holds for the lowest possible x i.e. $x = 0$, it also hold when x is approaching y . Because $f(0) = 0$, $f'(0) = f''(0) = 1$, by fixing $x = 0$, we can expand (83) as

$$1 + f'^2(y) - 2f'(y) < \left(2f'(y) + yf''(y) \right) f(y) + 2yf''(y) \quad (87)$$

which is equivalent to

$$\beta = \frac{1 + f'^2(y) - 2f'(y)}{(2f'(y) + yf''(y))f(y) + 2yf''(y)} < 1 \quad (88)$$

The general and complete analysis of changes in β is out of the scope of this paper and here I show it is robust on two class of convex functions that are defined on \mathbb{R}^+ .

If $f(x) = x^n + x^2/2 + x$ then,

$$f' = nx^{n-1} + x + 1, \quad f'' = n(n-1)x^{n-2} + 1 \quad (89)$$

$$(f' - 1)^2 = n^2x^{2n-2} + x^2 + 2nx^n \quad (90)$$

$$2f'f = 2\left(nx^{2n-1} + (1+n/2)x^{n+1} + (n+1)x^n + x^3/3 + 3x^2/2 + x \right) \quad (91)$$

$$yf''f = n(n-1)x^{2n-1} + (n(n-1) + 2)x^{n+1}/2 + n(n-1)x^n + x^3/2 + x^2 \quad (92)$$

and one can simply show that $(f' - 1)^2 < 2f'f + xf''f + 2xf''$. For a general polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + x^2/2 + x$, we have $\forall x > 1, \beta < 1$ (proof is possible by mathematical induction). Moreover, when the coefficients of small powers is not very larger than higher powers (e.g.

$\forall i < n, a_i \leq i$) then it holds in $x \in [0, 1]$ as well. Similarly for exponential function $f(x) = a^x / \ln(a)$, one can show that $\beta < 0$ always holds if $\ln(a) < 6$, otherwise it just holds for $x > 1$.

A.12 The optimal audit rate in self-assessment system

Similar to above proof in section A.6, in the absence of cross-checking, when e_b^s and e_c^s are jointly determined, define $e^s = e_b^s + e_c^s$, and from (22) and (23) assume

$$e^s = G\left(\frac{y_c^s}{\lambda_s}, y_c^s\right), \quad G_i = \frac{\partial G(x_1, x_2)}{\partial x_i}, \quad G_{ij} = \frac{\partial^2 G(x_1, x_2)}{\partial x_i \partial x_j} \quad (93)$$

then the condition (24) states that the optimal λ_s satisfies

$$\frac{\partial G}{\partial \lambda_s} = -\frac{y_c^s}{\lambda_s^2} G_1 = -\eta \quad (94)$$

Moreover, we can write

$$\frac{\partial^2 G}{\partial \lambda_s \partial y_c^s} = -\frac{1}{\lambda^2} G_1 - \frac{y_c^s}{\lambda^3} G_{11} - \frac{y_c^s}{\lambda^2} G_{12} \quad (95)$$

$$\frac{\partial^2 G}{\partial \lambda_s^2} = \frac{y_c^s}{\lambda_s} \left(\frac{2}{\lambda^2} G_1 + \frac{y_c^s}{\lambda^3} G_{11} \right) \quad (96)$$

Therefore, from (50) the derivative of λ_s w.r.t. y_c^s is obtained as

$$\frac{\partial \lambda_s}{\partial y_c^s} = \frac{\lambda_s}{y_c^s} \frac{1 - \gamma + \zeta}{2 - \gamma} \quad (97)$$

where $\gamma = \frac{y_c^s}{\lambda_s} \frac{G_{11}}{G_1}$ and $\zeta = y_c^s \frac{G_{12}}{G_1}$.

A.13 Comparing self-assessment and invoice cross-checking methods

The equivalent problem for self-assessment when cross-checking is possible is studied in Appendix A.1. In both cases the FOC w.r.t. e_c yields to (29). Therefore, from (30) we can write

$$\frac{\partial(e_b + e_c)}{\partial e_b} = 1 + \frac{\partial e_b}{\partial e_c} = \frac{\theta'(e_b + e_c) + \theta'(e_b)}{2\theta'(e_b + e_c) + e_c \theta''(e_b + e_c)} > 0 \quad (98)$$

One the other hand, the FOC w.r.t. e_b in the two problems yield to (28) and (22). Because of additional cost due to cross-checking, (28) gives smaller e_b than (22). Hence, using (98) total evasion is also smaller when cross-checking is available.

B Data appendix: VAT and final consumption in UK activities

VAT revenue is drawn from VAT Statistical Factsheets, October 2009, HM Revenue & Customs Office, UK: <http://www.uktradeinfo.com>. Other variables are drawn from Supply and Use Tables, 2004 - 2008, UK National Statistic, March 2011: <http://www.statistics.gov.uk>. All numbers are in £Millions. The industry classifications are not exactly the same in the two sources and some of the supply-use activities correspond to single VAT Factsheet code. Also, there are few cases that one supply-use activity corresponds to multiple VAT factsheet code. I narrowed the non-matching sectors down in each source, merged the single corresponding sectors together, and dropped the ones with zero or negative VAT payments (due to zero-rating of exports). Also, I dropped beverages and tobacco, since they are subject to excise duty.

Sector	VAT revenue	Gross value-added	Final consumption	Total demand
Forestry and logging	35	321	210	884
Manufacture of textiles	195	2609	14372	23538
Manufacture of wearing apparel	154	1385	33729	40067
Manufacture of leather and related products	26	491	7804	11221
Manufacture of wood, wood products and cork, except furniture	359	2671	1576	10633
Manufacture of paper and paper products	515	3445	4445	24638
Printing and reproduction of recorded media	654	16031	13947	37577
Manufacture of chemicals and chemical products	276	5109	2242	42110
Manufacture of rubber and rubber products	757	7175	4448	31105
Manufacture of other non-metallic mineral products	648	5173	3583	21398
Manufacture of fabricated metal products, except machinery & equipment	1430	13335	4276	41348
Manufacture of computer, electronic and optical products	457	2839	2408	33457
Manufacture of electrical equipment	372	13982	19877	122435
Manufacture of machinery and equipment n.e.c	113	12880	9682	62961
Manufacture of motor vehicles, trailers and semi-trailers	2295	9924	38602	97334
Manufacture of furniture	391	3954	11373	20223
Repair and installation of machinery and equipment	3767	27067	28307	45076
Electricity, gas, steam and air conditioning supply	442	16620	25032	66381
Sewerage	255	8388	9108	18918
Construction of buildings	4100	74619	7408	210196
Wholesale trade, except of motor vehicles and motorcycles	9286	46849	0	448
Retail trade except of motor vehicles and motorcycles	6303	62020	1753	2288
Postal and courier activities	333	8759	1013	15687
Telecommunications	2486	21806	15429	45809
Computer programming, consultancy and related activities	3649	35296	3	71168
Financial service activities, except insurance and pension funding	503	60083	38208	97837
Insurance and pension funding, except compulsory social security	453	17058	27251	52119
Activities auxiliary to financial services and insurance activities	177	13666	2347	29776
Real estate activities	1018	108713	113803	166621
Legal and accounting services	3811	30487	621	39243
Activities of head offices; management consultancy services	3120	15262	0	38033
Architectural and engineering activities; technical testing and analysis	2031	20610	327	40063
Scientific research and development	9	4947	327	12291
Advertising and market research	602	6491	82	23690
Other professional, scientific and technical activities	333	42198	1980	104594
Veterinary activities	138	65621	102655	114886
Rental and leasing activities	662	12104	7201	28376
Education	412	68926	77056	99472
Residential care activities	41	5273	5273	5285
Social work activities without accommodation	31	20344	32148	49576
Sports activities and amusement and recreation activities	1668	32348	39444	67386
Activities of membership organisations	213	7072	6053	10107